

EPO 59 104079

Europäisches  
PatentamtEuropean  
Patent OfficeOffice européen  
des brevets

REC'D 04 AUG 1999

WIPO PCT

Bescheinigung

Certificate

Attestation

Die angehefteten Unterla-  
gen stimmen mit der  
ursprünglich eingereichten  
Fassung der auf dem näch-  
sten Blatt bezeichneten  
europäischen Patentanmel-  
dung überein.

The attached documents  
are exact copies of the  
European patent application  
described on the following  
page, as originally filed.

Les documents fixés à  
cette attestation sont  
conformes à la version  
initialement déposée de  
la demande de brevet  
européen spécifiée à la  
page suivante.

Patentanmeldung Nr. Patent application No. Demande de brevet n°

98111506.6

**PRIORITY  
DOCUMENT**  
SUBMITTED OR TRANSMITTED IN  
COMPLIANCE WITH RULE 17.1(a) OR (b)

Der Präsident des Europäischen Patentamts;  
Im Auftrag

For the President of the European Patent Office

Le Président de l'Office européen des brevets  
p.o.

*Arlette Fiedler*

A. Fiedler

DEN HAAG, DEN  
THE HAGUE,  
LA HAYE, LE

27/07/99

11

11

11



Europäisches  
Patentamt

European  
Patent Office

Office européen  
des brevets

**Blatt 2 der Bescheinigung  
Sheet 2 of the certificate  
Page 2 de l'attestation**

Anmeldung Nr.:  
Application no.: 98111506.6  
Demande n°:

Anmeldetag:  
Date of filing: 23/06/98  
Date de dépôt:

Anmelder:  
Applicant(s):  
Demandeur(s):  
Biosearch Italia S.p.A.  
21040 Gerezano (VA)  
ITALY

Bezeichnung der Erfindung:  
Title of the invention:  
Titre de l'invention:

Methods for transferring the capability to produce a natural product from an original microorganism to a suitable production host

In Anspruch genommene Priorität(en) / Priority(ies) claimed / Priorité(s) revendiquée(s)

Staat:  
State:  
Pays:

Tag:  
Date:  
Date:

Aktenzeichen:  
File no.  
Numéro de dépôt:

Internationale Patentklassifikation:  
International Patent classification:  
Classification internationale des brevets:

/

Am Anmeldetag benannte Vertragsstaaten:  
Contracting states designated at date of filing: AT/BE/CH/CY/DE/DK/ES/FI/FR/GB/GR/IE/IT/LI/LU/MC/NL/PT/SE  
Etats contractants désignés lors du dépôt:

Bemerkungen:  
Remarks:  
Remarques:



BSI 0003 EPN

23.06.98

1

(4)

5                   **METHODS FOR TRANSFERRING THE CAPABILITY TO PRODUCE A  
NATURAL PRODUCT FROM A ORIGINAL MICROORGANISM TO A  
SUITABLE PRODUCTION HOST.**

**1. FIELD OF THE INVENTION**

10       The present invention relates to a novel approach for  
drug discovery. More particularly, the invention  
relates to a system for improving the process of lead  
optimization and development of compounds, when these  
compounds are natural products produced by  
15       microorganisms belonging to the order *Actinomycetales*  
or chemical derivatives of these compounds. The  
invention relates to a system for transferring the  
capability to produce a natural product from a  
microorganism belonging to the order *Actinomycetales*  
20       into a defined host, where said natural product can be  
optimally produced and its biosynthetic pathway  
suitably modified.

**2. BACKGROUND ART**

25       Natural products are complex molecules with important  
uses in medicine. Examples include: antibacterial  
agents, such as erythromycin, teicoplanin,  
tetracycline; antitumor compounds, such as  
dauxorubicin; antihelminthic compounds, such as  
30       ivermectin; immunosuppressive agents, such as  
cyclosporin and FK506; antifungal compounds, such as  
amphotericin and nystatin; etc. Natural products are  
produced as secondary metabolites by a wide range of  
living organisms. Although many secondary metabolites  
35       have been identified, there remains the need to obtain  
novel structures with new activities or enhanced  
properties. Current methods of obtaining such molecules

BSI 0003 EPN

11.23.06.98

2

include screening of natural isolates and chemical modification of existing ones. Random screening of natural products from disparate sources has resulted in the discovery of many important drugs and is still employed for seeking for novel activities. This process, which consists in exposing a miniaturized biological system to tens or hundreds of thousands of different compounds, in order to find those few that exhibit a desired property, is designated high throughput screening, or HTS.

One of the used sources widely in HTS is a collection of natural products produced by small-scale fermentation of newly isolated microorganisms. A natural product may have one or more potential therapeutic properties, including but not limited to antibacterial, antifungal, antiviral, antitumor, immunomodulating or other pharmacological properties. Natural products have long constituted a source of interesting, structurally original and "imaginative" molecules endowed with potent biological activities. In addition, recent observations indicate that only a small fraction of the microbial flora present in environmental samples, ranging from 0.01 to 1% according to the estimates, is related to known species. Microorganisms belonging to the order *Actinomycetales* represent thus far the group of producers unsurpassed for chemical and biological diversity. However, more than 15,000 natural products produced by microorganisms have been described, and the chances of finding new structures are relatively small, unless efforts are directed towards those classes of microorganisms that have been little exploited in the past. Poorly characterized actinomycete genera can thus constitute a useful source of novel structures. With proper methodologies, unusual genera can be isolated from environmental samples and some of these

BSI 0003 EPN

N 23 06 98

3

isolates will produce interesting activities. These could either represent completely new entities, or known molecules acting on a novel target or in a previously unreported way. Many of these products will

5 have original structures and potent biological activities. However, newly discovered secondary metabolites will be produced for the most part by microorganisms which have been isolated for the characteristic of being unusual and selected for their

10 ability to produce a given bioactivity. Consequently, little will be known about the best conditions for growth, productivity and storage. Often the microorganism does not produce a single bioactive compound, and other, unrelated activities must be

15 completely removed for a meaningful evaluation of the properties of the lead compound. Furthermore, rarely is a secondary metabolite produced as a single, bioactive molecule, but is often present as a "complex" of several, closely related compounds, only some of

20 which may possess the desired biological or chemical properties. Therefore, physiological conditions, such as nutrient and cofactor supply, that allow obtaining a "controlled" complex need to be established empirically by a trial and error approach. Finally, the natural

25 product may need to be structurally modified, and this can be achieved only by chemical modification. In essence, the scarce knowledge available on the physiology and genetics of the producing strain will severely hamper the lead optimization and development

30 processes.

Chemical modification of preexisting natural products has been successfully employed to generate derivatives of natural products, but it still suffers from practical limitations to the type of compounds

35 obtainable. Many natural products are often structurally complex molecules, with relatively large

BSI 0003 EPN

M 23 05 98

4

molecular weights. Due to their structural complexity, total synthesis of natural products is often prohibitive for the number of necessary steps and the overall yield; furthermore, selective modification of a natural product can often be efficiently performed only on limited portions of the molecule. This difficulty of generating structural derivatives by conventional medicinal chemistry slows down the process of lead optimization and supply. Microorganisms employ intricate biosynthetic machineries to make natural products: for example, synthesis of the macrolide antibiotic erythromycin, a secondary metabolite in the medium-range structural complexity, requires the participation of over 40 different enzymatic activities (Katz and Donadio, 1995, Macrolides, in Genetics and Biochemistry of Antibiotic Production, Vining and Stuttard eds., Butterworth-Heinemann, Boston CT, p. 385-420). Biosynthetic pathways can often be redirected through manipulation of the fermentation conditions or of the biosynthesis genes, in order to produce desired analogs of the original structure. The availability of genes involved in the formation of secondary metabolites has been exploited for the formation of derivatives of natural products obtained after genetic manipulation of the producing organism (Hopwood, 1997, Chem. Rev. 99:0-39). These manipulations have resulted in novel molecules, many of which would be extremely hard if not impossible to produce by chemical derivatization of the parent compound. The obvious economical and environmental benefits resulting from the formation of the desired structure in one fermentation step constitute an additional stimulus for the application of pathway engineering for the rational design of novel structures. The compounds obtained in this way are amenable evaluation of their biological properties as

BSI 0003 EPN

M 23 06 98

5

well as being substrates for further derivatization by chemical or biological means.

In summary, the supply of a natural product produced by a newly discovered microorganism, the optimization of the complex composition, and the process of lead optimization will all benefit from a detailed knowledge of the genetics and physiology of the producing strain. The present invention describes a general method for transferring the capability to produce any secondary metabolite from the original actinomycete producer to an established and genetically manipulatable production host. The general concept of the invention is illustrated in Fig. 1. Conditions for optimal growth, metabolite production and maintenance need therefore to be developed for one host. In addition, the availability of the cloned genes in a genetically manipulatable and well characterized host allows the utilization of all the genetic tools developed for these strains for the creation of novel derivatives of the natural product after genetic intervention.

### 3. SUMMARY OF THE INVENTION

The present invention provides a system for producing and manipulating natural products produced by a large group of bacteria for the purpose of drug discovery, development and production. The method of the invention transfers the ability to produce a secondary metabolite from an actinomycete that is the original producer of the natural product, to a another production host that has desirable characteristics.

In one embodiment, the invention involves the construction of a library from a donor organism, the producer of a natural product, in an artificial chromosome that can be shuttled between a convenient, neutral cloning host, such as the bacterium *Escherichia*

BSI 0003 EPN

M 23.06.98

6

coli, and a production host, such as the actinomycetes *Streptomyces lividans* or *Streptomyces coelicolor*. The clones directing the synthesis of the natural product are identified in said library, transferred into the production host where said natural product is synthesized.

In another embodiment, the invention involves the reconstruction of a large segment that directs the synthesis of a natural product, starting from smaller DNA fragments cloned from the genome of a donor organism. This reconstruction occurs in an artificial chromosome that can be transferred into an actinomycete production host and that is maintained in a convenient neutral host, such as the bacterium *Escherichia coli*. The reconstructed genomic segment in the artificial chromosome is transferred into the production host where said natural product is synthesized.

The present invention also relates to *Escherichia coli*-*Streptomyces* Artificial Chromosomes, recombinant DNA vectors useful for shuttling the genetic information necessary to synthesize a given natural product between a donor actinomycete producer and a production host.

### 3.1 DEFINITIONS

As used herein, the following terms will have the meaning indicated.

An "*Escherichia coli*-*Streptomyces* Artificial Chromosome" is a recombinant DNA vector that can accept and maintain very large DNA inserts in an *Escherichia coli* host, and that can be introduced and maintained in an actinomycete production host.

A "natural product" is a secondary metabolite made by a microorganism through a series of biosynthetic steps. This natural product may or may not have any useful biological activity.

BSI 0003 EPN

7

A "complex" is the mixture of related natural products with similar properties and biological activity that are often produced by the same biosynthetic pathway.

- 5 A "donor organism" is the original producer of a natural product, where the synthesis of said compound is governed by a defined number of genetic elements.

- A "gene cluster", a "cluster", a "biosynthesis cluster" all designate a contiguous segment of the  
10 donor organism's genome that contains all the genes required for the synthesis of a natural product.

- A "producing host" is a microorganism where the production of a natural product is directed by a gene cluster derived from a donor organism and introduced  
15 into the production host via an *Escherichia coli*-*Streptomyces* Artificial Chromosome.

- As used in the present invention, the following abbreviations are employed: °C (Celsius degree); h (hour); min (minute); kb (kilobase); µl (microliter);  
20 ml (milliliter); mm (millimeter); mg (milligram); µg (microgram); ng (nanogram); M (molar); Mb (megabase); UV (ultraviolet); kV (kilovolt); Ω (Ohm); mFa (millifaraday).

- In addition, the following abbreviations are used:  
25 Ab, antibiotic; Ap, ampicillin; bp, base pair; ca., circa (i.e. "about"); Cm, chloramphenicol; ESAC, *Escherichia coli*-*Streptomyces* Artificial Chromosome; *E. coli*, *Escherichia coli*; GC, guanosine + cytosine; HTS, high throughput screening; Km, kanamycin; LB, Luria  
30 Broth; LMP, low melting point; NMR, nuclear magnetic resonance; MS, mass spectrometry; PCR, polymerase chain reaction; PFGE, Pulsed Field Gel Electrophoresis; *P. rosea*, *Planobispora rosea*; <sup>R</sup>, resistance; rpm, rounds per minute; <sup>S</sup>, sensitive; SDS, sodium dodecyl sulfate;

BSI 0003 EPN

M 23.06.98

8

*S. coelicolor*, *Streptomyces coelicolor*; *S. hygroscopicus*, *Streptomyces hygroscopicus*; *S. lividans*, *Streptomyces lividans*; Tc, tetracycline; TE, TrisHCl EDTA buffer; Th, thiostrepton; ts, temperature sensitive; U, units; vol, volume; wt, weight; YEME, yeast extract malt extract medium.

#### 4. BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1. Scheme of the invention. The general concept of the invention, whereby the gene cluster required for the synthesis of a natural product in a donor organism is established in an ESAC vector in an *Escherichia coli* host, and then transferred into a desired production host, where they integrates into the chromosome and directs production of the secondary metabolite. The hexagon represents the natural product, the twisted thin line the bacterial chromosomes, and the thick line the desired gene cluster. The ESAC episome is represented by a circle.

Figure 2. *E. coli*-*Streptomyces* artificial chromosome vectors. Vectors pPAC-S1 and pPAC-S2 differ solely for the orientation of the *int*-*tsr* cassette. Relevant features of the vectors are illustrated. Suitable cloning sites are shown as: B, *Bam*HI; S, *Sca*I; X, *Xba*I. The replicating function of bacteriophage P1 are indicated by the thick bars.

Figure 3. General scheme of the invention, top-down approach. High molecular weight DNA from the donor organism is cloned into an ESAC vector. The resulting library in *E. coli* is screened with the required probes, and the relevant ESAC clones are identified. These are introduced into the desired production host strain, where they integrate site-specifically into the host chromosome. Symbols and abbreviations are as in Fig. 1.

Figure 4. General scheme of the invention,

BSI 0003 EPN

9

bottom-up approach. A cosmid library is prepared with DNA from the donor organism and screened with the required probes. The overlapping inserts from the positive cosmids are assembled into an ESAC vector via homologous recombination in *E. coli*. The reconstructed ESAC clone is introduced into the desired production host, where it integrates site-specifically into the host chromosome. Symbols and abbreviations are as in Fig. 1.

Figure 5. Scheme of assemblage. The figure illustrates a hypothetical genomic segment from a donor organism that is covered by the inserts from three overlapping clones. The relevant fragments A and D, which denote the ends of the segment, and B and C, which represent regions of overlap, are indicated with their relative orientation (thick side on the fragment rectangle). The bottom part illustrates the reconstructed ESAC clone.

Figure 6. Constructs required for cluster assemblage. The plasmids indicated are generated by routine in vitro DNA manipulations. Fragments A, B, C and D are as in Fig. 5. Fragment pairs are in this example separated by a marker, indicated as  $Ab^R$  for antibiotic resistance. Selective markers present on the two compatible replicons are, as an example:  $Cm^R$  and  $Km^R$ .

Figure 7. Interplasmid insert exchange. Each of the  $Cm^R$  derivatives, as of Fig. 6, is introduced in the same *E. coli* cell as the cognate clone of Fig. 5 (for example a cosmid that carries a  $Km^R$  marker). Formation and then resolution of the cointegrate leads to the transfer of the cosmid's insert, indicated here by a looping L, in the  $Cm^R$  replicon.

Figure 8. Sequel of assembling steps. A series of interplasmid cointegration and resolution events is conducted. Only the growing ESAC clone is indicated.

BSI 0003 EPN

23.06.98

10

The starting ESAC clone (Fig. 6) is recombined with plasmid pAB2 (Fig. 7), leading to the insertion into ESAC of the insert flanked by fragments A and B. Next, the  $Ab^R$  from pBC1 (Fig. 6) is introduced between  
5 fragments B and C, and subsequently replaced by the insert from pBC2 (Fig. 7). Finally, the  $Ab^R$  from pCD1 (Fig. 6) is introduced between fragments C and D, and subsequently replaced by the insert from pCD2 (Fig. 7).

Figure 9. The GE2270 cluster. A restriction map  
10 of the *Bam*HI sites (indicated as short vertical lines) in the GE2270 gene cluster from *Planobispora rosea* ATCC 53733 is reported, together with the cosmids pRP16, pRP31 and pRP58. The fragments A, B, C and D used for assemblage are highlighted. Restriction sites are  
15 abbreviated as: M, *Sma*I; P, *Pst*I; S, *Sac*I.

Figure 10. Signature sequences at the left (panel A) and right (panel B) ends of the GE2270 cluster. The sequence in panels A starts around coordinate 1.8 kb (Fig. 9); the sequence in panel B ends around  
20 coordinate 91.0 kb (Fig. 9). The orientation of the sequences is the same as in Fig. 9.

Figure 11. Site-specific integration. PFGE analysis of *S. lividans* ZX7 transformed with ESAC-70. Lanes 1 and 2: *S. coelicolor* M145; lane 3: *S. lividans*  
25 ZX7 DNA; lane 4: ZX7 *attB*::ESAC-70 DNA, colony 1; lane 5: ZX7 *attB*::ESAC-70 DNA, colony 2; lane 6: 50-kb ladder, size marker. All DNAs in lanes 1-5 are digested with *Dra*I. Conditions for PFGE are: 200 Volts, 70 s switching for 7 15 h, 120 s switching for 11 h.

30

## 5. DETAILED DESCRIPTION OF THE INVENTION

In its broadest sense, the present invention entails a general procedure for constructing a *Streptomyces* host producing any natural product after selective transfer  
35 of the relevant genes from the original actinomycete producer, the donor strain. This general procedure is

BSI 0003 EPN

11

outlined in Fig. 1. The present invention can be applied with only limited information on the structure of the natural product and very little knowledge of the original producer's genetics. The present invention

5 has a substantial impact on the process of drug discovery involving natural products or their structural derivatives. The transfer of the producing capability to a better characterized host can substantially improve several portions of the process

10 of lead optimization and development: the titer of the natural product in the producing strain can be more effectively increased; the purification of the natural product can be carried out in a known background of possible interfering activities; the composition of the

15 complex can be more effectively controlled; altered derivatives of the natural product can be more effectively produced through manipulation of the fermentation conditions or by pathway engineering. In order to better understand the value of the present

20 invention, a brief description is reported below of the current methods for optimizing the productivity of the producing strain, for purifying a natural product, for controlling the composition of a complex, and for producing derivatives of a natural product.

25 The production of a natural product is controlled by several mechanisms, few of which have been established in detail. Generally, the level of production of a natural product depends on the composition of the growth medium; on the presence of

30 appropriate precursors or on the absence of specific inhibitors; on the timing and level of expression of genes of the biosynthetic pathway and of competing routes; and on the level and specific activity of key enzymes in the pathway. Because of this complexity, the

35 productivity of the original strain is usually increased by an empirical process, which may include,

BSI 0003 EPN

M 23 08 98

12

among other things, one or more of the following steps:  
strain purification, selection of phenotypic variants  
arising spontaneously or after mutagenic treatment of  
the strain, variation in the fermentation medium or in  
5 the fermentation parameters; genetic engineering of the  
producing strain. Fundamental knowledge about the  
physiology of the producing strain and the variables  
affecting titer must be achieved for an effective  
improvement of productivity. This knowledge is very  
10 scant in a newly identified producer strain.

During the discovery and development phase,  
sufficient quantities of a natural product must be  
available for an evaluation of its properties and/or  
for the generation of analogs. Because of its  
15 uniqueness, a specific purification process must be  
developed for each natural product. However, it is  
highly desirable to have the natural product as free as  
possible of compounds that may interfere with the  
biological activity of the molecule. Contaminating  
20 impurities must be characterized analytically and  
biologically. In a poorly characterized producer,  
little information is available on the relevance of  
contaminating impurities.

A natural product may be produced by a  
25 microorganism as a complex of a few or tens of  
molecules with minor structural differences, designated  
congeners. Although most of the congeners are usually  
biologically active, only one or a few may represent  
the desired product: for example, one congener may be  
30 substantially more active than the others; it may  
possess better physico-chemical properties; or it may  
be a better substrate for chemical modification. The  
composition of a complex can be somehow controlled by  
intervening on the fermentation parameters. However,  
35 the most effective way is usually the altered  
expression of selected genes by genetic engineering

BSI 0003 EPN

N 23.08.98

13

(e.g. Sezonov et al., 1997, Nature Biotechnol. 15:349-353).

Chemical modification of natural products represents the most commonly used means of obtaining novel structures. This approach has been successfully employed, but it still suffers from practical limitations to the number and type of compounds obtainable. The structural complexity of many natural products makes their total synthesis often too lengthy and expensive to be of any practical use. This same structural complexity, with either the presence of several closely related functional groups or their absence, limits modification of a natural product to selected portions of the molecule. Methods of combinatorial synthesis need an initial scaffold as the starting building block, and this can be often generated only through a low yield degradation of the natural product. However, derivatives of natural products that would be very hard if not impossible to produce by chemical means have been obtained after genetic alteration of the biosynthetic pathway. Examples include the introduction of additional genetic information (Epp et al., 1989, Gene 85:293-301), the targeted inactivation of selected genes or portion thereof (Donadio et al., 1993, Proc. Natl. Acad. Sci. USA 90:7119-7123), the "mixing and matching" of genes or portions thereof from different pathways (McDaniel et al., 1994, Nature 375:549-554).

All the above activities are important for the process of lead optimization and for the development of selected lead structures. They can all benefit, to different extent, from a detailed knowledge of the physiology of the producing strain, and from the possibility of genetically manipulating it. The process by which a given organism is genetically manipulated in order to alter the type, quality or quantity of a

BSI 0003 EPN

23-06-98

14

natural product is referred to as pathway engineering.

The ability to perform pathway engineering in a newly isolated microorganism producing a bioactive molecule with promising characteristics can therefore

5 considerably expedite the optimization of a lead structure and the development process. Pathway engineering can be schematized as a sequel of three steps: a) isolation of the genes of interest; b) performing on selected gene(s) the manipulations  
10 required by the specific objective; and c) introduction of the modified gene(s) in suitable form in an appropriate host.

Isolation of the genes of interest from most actinomycetes can be achieved quite easily. The genes  
15 for primary metabolism are usually well conserved, and they can be easily accessed in any microorganism by using suitable hybridization probes or by the PCR. The genetic elements governing the biosynthesis of the major classes of secondary metabolites have been also  
20 described, and many genes can similarly be identified.

Since natural product biosynthesis is governed by clusters, one needs to identify just a few genes in order to have them all. However, synthesis of the vast majority of natural products requires a considerable  
25 extent of genetic information. For examples, biosynthesis of the natural products erythromycin (an antibiotic), avermectin (an antihelmintic agent) and rapamycin (an immunosuppressant) requires 55, 90 and 95 kb, respectively, of genetic information (Katz and  
30 Donadio, 1993, *Annu. Rev. Microbiol.* 47:875-912; MacNeil, 1995, *Avermectins*, in *Genetics and Biochemistry of Antibiotic Production*, Vining and Stuttard eds., Butterworth-Heinemann, Boston CT, p.421-442; Schwecke et al., 1995, *Proc. Natl. Acad. Sci. USA*  
35 92:7839-7843). Other natural products may require even larger extent of genetic information. Therefore, in

BSI 0003 EPN

N 23 06 98

15

order to isolate an entire cluster in a single piece, cloning vectors capable of accepting and maintaining large DNA segments are necessary.

The manipulation of the isolated genes is  
5 generally best performed in a convenient cloning host, such as *E. coli*. Manipulations relevant to pathway engineering can include some or all of the following: site directed mutagenesis, gene inactivation, gene fusions, modification of regulatory sequences, etc.  
10 Techniques for the in vitro manipulation of DNA and for the propagation of the mutated alleles in *E. coli* are very well developed and can be applied to DNA from virtually any source (Sambrook et al., 1989, In  
15 *Molecular Cloning: A laboratory Manual*, 2nd edn, Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press).

The final step in pathway engineering requires the introduction of modified or heterologous gene(s), in suitable form, in a strain where these genes can be  
20 appropriately expressed. This strain is often the strain producing the natural product whose quantity, quality or type one wants to alter. The genes of interest must be carried on appropriate vectors: according to the particular objective of pathway  
25 engineering, one may need, among others, vectors that can be stably maintained as single or multicopy episomes; that can insert into the host chromosome at a fixed location; that allow replacement of an endogenous gene with an *in vitro* modified allele; that allow  
30 deletion of selected genes from the host chromosome. In addition, for each strain one must have means for introducing heterologous DNA and for selecting for its presence. Therefore, in order to genetically manipulate a given producer, one must establish conditions for  
35 rendering the bacterial cell capable of receiving incoming DNA; for selecting the incoming DNA; and

BSI 0003 EPN

M 23 05 98

16

develop vectors and methodologies for the various types of manipulations exemplified above. Low- and high copy-number, integrative, non-replicating vectors must also be developed with appropriate selection markers.

5 Therefore, for each producing strain, specific gene transfer tools and conditions must be developed, starting in most cases from extremely poor knowledge about the microorganism. In addition, techniques developed for one species do not necessarily apply to a  
10 new species from the same genus, and often not even to a new strain. It is then no wonder that, among the thousands of strains described as producers of interesting natural products, gene transfer systems have been developed only for a limited number of  
15 species, which serve either as model organisms for genetic and physiological studies, or produce a commercially important molecule. The present invention provides tools for the general manipulation of any secondary metabolite pathway, and overcomes the  
20 difficulty of developing *ad hoc* conditions for a new producer.

Naive hosts have been shown to produce the appropriate natural product or its intermediate(s) when the relevant DNA was introduced into them (Malpartida  
25 and Hopwood, 1984, *Nature* 309:462-464; Hong et al., 1997, *J. Bacteriol.* 179:470-476; Kao et al., 1994, *Science* 265:509-512; McGowan et al., 1996, *Mol. Microbiol.* 22:415-426; Kealey et al., 1998; *Proc. Natl. Acad. Sci. USA* 95:505-509). However, the examples  
30 reported thus far have represented special cases. Indeed, they have reported the introduction of relatively small DNA segments into a production host; or the transfer of gene clusters within members of the same bacterial genus; or they required the careful  
35 engineering of specific biosynthesis genes under the control of appropriate genetic elements that would

BSI 0003 EPN

M 23 06 98

17

direct their expression. The *Streptomyces* vectors currently available have an upper limit of ca. 40 kb (Hopwood et al., 1987, Methods Enzymol. 153:116-167).

5 The unexpected finding described in the present invention is that all the genetic elements required for the synthesis of a natural product in the original producer can be conveniently expressed in a heterologous host, where they direct the synthesis of the desired molecule. It was also unexpected and  
10 unprecedented that this heterologous expression can occur when the donor organism and the production host belong to different bacterial genera. Furthermore, up until now, it was not established that DNA fragments exceeding 100 kb, derived from the high GC genome of  
15 actinomycetes, could be cloned and stably maintained in an *E. coli* host. Nor was any report of the introduction of large DNA segments into a *Streptomyces* host.

The present invention rests on the fact that the genes required for the formation of a natural product  
20 are found as gene clusters of a defined size; that these gene clusters can be conveniently isolated, manipulated and transferred among different actinomycete strains; that they will be expressed in a heterologous host; and on the fact that all the primary  
25 metabolite precursors required for the formation of a particular natural product are either produced by selected enzymes encoded by cluster-specific genes, or are present and available in the heterologous host at the time of formation of the natural product. The  
30 present invention addresses also the crucial aspect of natural product formation in actinomycetes: i.e. synthesis of many natural products requires over 100 kb of genetic information. To be generally applicable, transferring all the genes necessary for the production  
35 of any natural product requires cloning vectors capable of accommodating fragments as large as 150 kb, and

BSI 0003 EPN

M 23.06.98

18

possibly more. An object of the present invention is therefore represented by vectors capable of accommodating such large fragments which are also capable of being stably maintained in a suitable microbial host, such as a *Streptomyces* host.

Examples of these vectors are designated with the generic name ESAC: *E. coli-Streptomyces* Artificial Chromosomes. They are derived from bacterial artificial chromosomes (Shizuya et al., 1992, Proc. Natl. Acad. Sci. USA 89:8794-8797; Ioannou et al., 1994, Nature Genet. 6:84-89), which can carry inserts up to 300 kb, or more.

As a general example of the broad applicability of the principles and methodologies described in the present invention, the Examples reported below describe how a convenient *Streptomyces* host can be engineered to produce a desired natural product after mobilization of the corresponding gene cluster through the use of an appropriate ESAC vector. The exemplary organism chosen as the original producer of the natural product is the actinomycete *P. rosea*, belonging to one of the lesser characterized genera of actinomycetes (Goodfellow, 1992, In *The Prokaryotes*, 2nd edn., Balows, Trueper, Dworkin, Harder and Schleifer eds, Springer-Verlag, New York, NY, USA). This organism produces the natural product GE2270 (Selva et al., 1991, J. Antibiotics 44:693-701), an antibacterial agent. However, little is known about the biosynthesis of this molecule, nor has any biosynthesis gene been described. This particular case therefore describes the general applicability of the present invention, since very little information is available on the donor organism, on its genetics and physiology, and on the gene cluster responsible for the biosynthesis of the natural product GE2270. In order to transfer the capability to produce GE2270 in a convenient production host, it is necessary to identify

BSI 0003 EPN

M 23.05.98

19

the GE2270 gene cluster. This is achieved by the use of selected hybridization probes, as described in detail in the Examples reported below.

The present invention, relating to a general  
5 methods for transferring the capability to produce any natural product from the original actinomycete to an established and genetically manipulatable *Streptomyces* host, can be schematized in a series of passages summarized as: 1) design of suitable vectors; 2)  
10 construction of a large-insert library in said vectors; 3) selection of the desired clones with appropriate probes; 4) insertion of the selected clones into a convenient *Streptomyces* host; and 5) growth of the recombinant strain under appropriate conditions to  
15 produce the natural product.

Actinomycetes produce a large number of natural products with important applications. However, other important classes of microbial producers are known, and newer ones are likely to be discovered in the upcoming  
20 years, as more microbial sources are screened for potential new drugs. Important classes of microbial producers include, among others, filamentous fungi, mixobacteria, pseudomonas and cyanobacteria. The series of passages described above can therefore be  
25 applied to other important classes of microbial producers, provided that two requisites are met; the synthesis of the desired natural product is governed by a gene cluster; suitable production host(s) exist; and appropriate selective marker(s) and maintenance  
30 function(s) are introduced into the artificial chromosome.

Therefore, the principles and methodologies described in the present invention can be extended to microbial producers other than actinomycetes.

35 Furthermore, the series of passages summarized above and described in detail in the Examples, involve

M 23.06.98

BSI 0003 EPN

20

the use of a neutral, cloning host. This host, as described in the present invention, is the bacterium *Escherichia coli*. A preferred example of this hosts, a high cloning efficiency can be obtained, an many of the analyses of the ESAC clones can be quickly performed. However, it is evident to one or ordinary skill in this art that any other host that allows high cloning efficiency can be used as neutral, cloning host. Additionally, the use of a neutral cloning host is not a *conditio sine qua non* for the applicability of the present invention. In fact, when it is possible to establish directly a library in a production host, there is no need for an intermediate neutral cloning host.

15

## 6. GENERAL METHODS

### Plasmids, Bacterial Strains and Growth Conditions.

Plasmids pUCBM20, pUCBM21, pBR322 and pUC18 are obtained from Boheringer Mannheim; plasmid pIJ39 and  $\Phi$ C31 DNA are from prof. David Hopwood, The John Innes Centre, Norwich, UK; plasmid pCYPAC2 is from prof. Pieter de Jong, Roswell Park Cancer Institute, Buffalo, NY, USA; plasmid pMAK705 from prof. Sidney Kushner, University of Georgia, Athens, USA; cosmid Lorist6 from prof. Stewart Cole, Pasteur Institute, Paris, France. *E. coli* strains DH5 $\alpha$ , DH10B, C600, DH1 and XL1blue are obtained from commercial sources. *S. coelicolor* M145 and *S. lividans* ZX7 are from prof. David Hopwood, The John Innes Institute, Norwich, UK. *Planobispora rosea* ATCC 53733 and *Streptomyces hygroscopicus* ATCC 29253 are from the ATCC culture collection. All other materials are from commercial sources. Media for cultivation of *E. coli* (Sambrook et al., 1989, In *Molecular Cloning: A laboratory Manual*, 2nd edn, Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press) and *Streptomyces* (Hopwood et al., 1985, Genetic

BSI 0003 EPN

21

Manipulation of *Streptomyces*: A Laboratory Manual, The John Innes Foundation, Norwich, UK) have been described. The JM medium for *S. coelicolor* has been described (Puglia et al., 1995, Mol. Microbiol. 17:737-746).

DNA Manipulations DNA manipulations are performed following described procedures, using the appropriate *E. coli* strains as cloning hosts (Sambrook et al., 1989, In *Molecular Cloning: A laboratory Manual*, 2nd edn, Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press). Amplifications by the PCR Genomic DNA from actinomycetes is prepared as described (Hopwood et al., 1985, *Genetic Manipulation of Streptomyces: A Laboratory Manual*, The John Innes Foundation, Norwich, UK). A cosmid library of *P. rosea* DNA is constructed in the cosmid vector Lorist6 following published procedures (Sambrook et al., 1989, In *Molecular Cloning: A laboratory Manual*, 2nd edn, Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press). Amplification by the PCR are performed following published guidelines (Innis, Gelfand, Sninsky and White, eds., 1990, *PCR Protocols: A guide to Methods and Applications*, Academic Press, San Diego, CA, USA).

Hybridizations Probes Pep6 and Pep8 are derived from conserved motifs in peptide synthetase gene sequences (Turgay and Marahiel, 1994, Pept. Res. 7:238-241). Oligonucleotide probe Pep6 consists of an equimolar mixture of oligonucleotides 5'-GCSTACATCATCTACACSTCSGGSACSACSGGSAAGCCSAAGGG-3' and 5'-GGSTACATCATCTACACSAGCGGSACSACSGGSAAGCCSAAGGG-3'. Oligonucleotide probe Pep8 consists of an equimolar mixture of oligonucleotides 5'-AKGCTGTCSCCSCCSAGSNNGAAGAAGTYGTCGTCGATSCC-3' and 5'-AKGGAGTCSCCSCCSAGSNNGAAGAAGTYGTCGTCGATSCC-3'. [S indicates G or C; K indicates G or T; Y, C or T; and N,

BSI 0003 EPN

M 23.05.98

22

any base]. Hybridizations are performed with a hybridization stringency set at 2xSSC, 55 °C, and a final wash set at the same stringency.

Preparation of high molecular weight DNA

- 5 Procedures for the preparation of high molecular weight DNA from actinomycetes for PFGE have been described (Dyson, 1993, Trends Genet. 9:72; Kieser et al., 1992, J. Bacteriol. 174:5496-5507). They are slightly modified for constructing libraries as described in the
- 10 Examples.

- Metabolite characterization GE2270 is recovered as previously described (Selva et al., 1991, J. Antibiotics 44:693-701). Structural determination is performed according to published procedures (Tavecchia
- 15 et al., 1995, Tetrahedron 51:4867-4890).

**7. EXAMPLES**

- The present invention consists in a series of passages, involving the design of suitable vectors; the
- 20 construction of a large-insert library in said vectors employing genomic DNA from the donor organism; the selection of the desired clones carrying the cluster specifying the synthesis of the desired natural product; the introduction of selected clone(s) into the
- 25 appropriate production host; and the production of the natural product by the recombinant strain under appropriate conditions. These passages are described in detail in the Examples reported herein. These Examples outline the steps necessary to accomplish each passage,
- 30 for the overall purpose of the present invention: the production of a natural product in a different host. The Examples serve to illustrate the principles and methodologies of the present invention, and are not meant to restrict its scope to the Examples specified
- 35 herein.

BSI 0003 EPN

23

### 7.1 Cloning vectors

Bacterial artificial chromosomes are circular plasmids that can be easily propagated in and prepared from *E. coli* cells by standard miniprep methods (Shizuya et al., 1992, Proc. Natl. Acad. Sci. USA 89:8794-8797; Ioannou et al., 1994, Nature Genet. 6:84-89). In order to adapt bacterial artificial chromosomes to a *Streptomyces* host, they need to be endowed with a selectable marker and maintenance functions. Site-specific integration, mediated by the action of an integrase encoded by the *int* gene, allows the stable incorporation of episomal elements into the host genome, at a defined locus designated *attB*. The episomal element needs to carry the cognate *attP* site and it may lack replicative functions. In addition, *int*-mediated excision of the integrated element from the chromosome via reversal of the integration event can be prevented through selection of the resistance marker carried by the integrated episome or, if necessary, after site-specific integration has occurred, the *int* gene on the integrated episome can be inactivated. Site-specific integration therefore allows the introduction of foreign DNA in single copy at a defined genetic locus. Several systems capable of directing site-specific integration of incoming circular DNA into the chromosome of a *Streptomyces* host have been described. A convenient system that can be used in the present invention is for instance the *int-attP* system derived from the temperate bacteriophage  $\Phi$ C31 (Kuhstoss and Rao, 1991, J. Mol. Biol. 222:897-908). The *int-attP* system of  $\Phi$ C31 naturally directs integration of the 41-kb phage genome during lysogen formation. The *attB* site in *S. coelicolor* is located in a stable segment of the chromosome (Redenbach et al., 1985, Mol. Microbiol. 21:77-96). Several selectable markers have been described that can be used for

BSI 0003 EPN

M 23.06.98

24

*Streptomyces* (Hopwood et al., 1985, Genetic Manipulation of *Streptomyces*: A Laboratory Manual, The John Innes Foundation, Norwich, UK). The *tsr* gene, conferring resistance to the antibiotic thiostrepton

5 (Thompson et al., 1982, Gene 20:51-62), is used in the present invention. The ESAC vectors pPAC-S1 and pPAC-S2, described in the present invention, are depicted in Fig. 2. Their relevant features are: ability to accommodate DNA inserts up to 300 kb; low copy number

10 in *E. coli* for increased stability; ease of propagation in *E. coli* because of the pUC19 stuffer segment; *Bam*HI, *Xba*I or *Sca*I cloning sites, with positive selection of inserts for resistance to sucrose; T7 and SP6 promoters flanking the cloning site; Km<sup>R</sup> or Th<sup>R</sup> for selection in

15 *E. coli* or actinomycetes, respectively; site-specific integration at the  $\Phi$ C31 *attB* site into the *Streptomyces* genome. Vectors pPAC-S1 and pPAC-S2 are 22 kb in size and differ solely for the orientation of the *int*-*tsr* cassette. After release of the stuffer pUC19

20 segment, the vector size is reduced to 19.7 kb. When cloning in the *Bam*HI site, the vector can be released by digestion with *Dra*I, resulting in vector fragments of 7.4, 4.2 and 0.6 kb. The additional 7.5 kb of vector DNA will be associated with the insert. *Dra*I rarely

25 cuts in the high-GC genome of actinomycetes, so that the insert size can be easily calculated.

#### Example 1

##### Isolation of the *int* region from $\Phi$ C31

30 Two pairs of PCR primers, 5'-TTTTTGGTACCTGACGTCCCGAAGGCGTG-3' and 5'-CAGCTTGTCATGGCGGA-3'; and 5'-TCTGTCCGCCATGGACAAGC-3' and 5'-TTTTTGGATCCGGCTAACTAACTAAACCGAGA-3', are used to amplify the *int*-containing fragments of 1.3 and 0.9

35 kb, respectively. The template is  $\Phi$ C31 DNA. The amplified fragments are digested with *Kpn*I + *Nco*I and

BSI 0003 EPN

25

*Nco*I + *Bam*HI, respectively, and recovered from an agarose gel.

#### Example 2

##### 5                    Construction of plasmid pINT

The 1.3 and 0.9 kb fragment, prepared as described in Example 1, are ligated to pUCMB21, digested with *Kpn*I + *Bam*HI. The resulting mixture contains the desired plasmid pINT.

10

#### Example 3

##### Construction of *E. coli* K12 DH5 $\alpha$ /pINT

Approximately 10 ng of plasmid pINT, prepared as described in Example 2, are used to transform *E. coli* DH5 $\alpha$  and a few of the resulting Ap<sup>R</sup> colonies that appear on the LB-agar plates are analyzed for their plasmid content. One colony is found to carry pINT, as verified by the observation, upon agarose gel-electrophoresis, of fragments of 4.0 and 0.9 kb after digestion of the plasmid with *Nco*I + *Bam*HI; and of 3.6 and 1.3 kb after digestion with *Kpn*I + *Nco*I.

20

#### Example 4

##### Construction of plasmids pUIT1 and pUIT2

25 The 1.8 kb *Bam*HI fragment containing the *tsr* gene is isolated from pIJ39 and ligated to pINT, prepared as described in Example 3, and previously digested with *Bam*HI. The resulting mixture contains the desired plasmids pUIT1 and pUIT2.

30

#### Example 5

##### Construction of *E. coli* K12 DH5 $\alpha$ /pUIT1 and DH5 $\alpha$ /pUIT2

Approximately 10 ng of plasmid pUIT1 and pUIT2, prepared as described in Example 4, are used to transform *E. coli* DH5 $\alpha$  and a few of the resulting Ap<sup>R</sup>

35

N 23.06.98

BSI 0003 EPN

26

colonies that appear on the LB-agar plates are analyzed for their plasmid content. One colony is found to carry pUIT1, as verified by the observation, upon agarose gel-electrophoresis, of fragments of 4.9 and 1.8 kb after *Bam*HI digestion of the plasmid; and of 3.0, 3.0 and 0.7 kb after *Eco*RI + *Sac*II digestion. Another colony is found to carry pUIT2, as verified by the observation, upon agarose gel-electrophoresis, of fragments of 4.9 and 1.8 kb after *Bam*HI digestion of the plasmid; and of 3.0, 2.7 and 1.0 kb after *Eco*RI + *Sac*II digestion.

#### Example 6

##### Construction of plasmid pUIT3

The 3.7 kb *Apa*I fragment, containing the *int*-*tsr* cassette, is isolated from plasmid pUIT1, prepared as described in Example 5, and ligated to pUCBM21 digested with *Apa*I. The resulting mixture contains the desired plasmid pUIT3.

20

#### Example 7

##### Construction of *E. coli* K12 DH5 $\alpha$ /pUIT3

Approximately 10 ng of plasmid pUIT3, prepared as described in Example 6, are used to transform *E. coli* DH5 $\alpha$  and a few of the resulting Ap<sup>R</sup> colonies that appear on the LB-agar plates are analyzed for their plasmid content. One colony is found to carry pUIT3, as verified by the observation, upon agarose gel-electrophoresis, of fragments of 3.7 and 2.7 kb after *Apa*I digestion of the plasmid; and of 4.2 and 2.2 kb after *Bam*HI digestion.

30

#### Example 8

##### Construction of plasmid pUIT4

The *Bam*HI site present in the *int*-*tsr* cassette of plasmid pUIT3 is eliminated as follows. Plasmid pUIT3,

35

BSI 0003 EPN

27

prepared as described in Example 5, is partially digested with *Bam*HI, followed by filling-in of the resulting ends, and treated with DNA ligase. The resulting mixture contains the desired plasmid pUIT4.

5

#### Example 9

##### Construction of *E. coli* K12 DH5 $\alpha$ /pUIT4

Approximately 10 ng of plasmid pUIT4, prepared as described in Example 8, are used to transform *E. coli* DH5 $\alpha$  and a few of the resulting Ap<sup>R</sup> colonies that appear on the LB-agar plates are analyzed for their plasmid content. One colony is found to carry pUIT4, as verified by the observation, upon agarose gel-electrophoresis, of a 6.4 kb fragment, and of 3.1, 2.9 and 0.2 kb fragments after *Bam*HI and *Nru*I digestion, respectively.

15

#### Example 10

##### Construction of plasmid pPAC-S1 and pPAC-S2

The 3.7 kb *Apa*I fragment from pUIT4, prepared as described in Example 9, is mixed with pCYPAC2, previously digested with *Nhe*I. After filling-in of the ends, DNA ligase is added. The resulting mixture contains the desired plasmids pPAC-S1 and pPAC-S2.

25

#### Example 11

##### Construction of *E. coli* K12 DH10B/pPAC-S1 and DH10B/pPAC-S2

Approximately 10 ng of plasmids pPAC-S1 and pPAC-S2, prepared as described in Example 10, are used to transform *E. coli* DH10B and a few of the resulting Km<sup>R</sup> colonies that appear on the LB-agar plates are analyzed for their plasmid content. One colony is found to carry pPAC-S1, as verified by the observation, upon agarose gel-electrophoresis, of fragments of 8.1, 4.8, 4.6, 2.2, 2.2, 0.5 and 0.1 kb after *Eco*RI digestion of the

35

BSI 0003 EPN

M 23.06.98

28

plasmid. Another colony is found to carry pPAC-S2, as verified by the observation, upon agarose gel-electrophoresis, of fragments of 8.1, 7.8, 2.2, 2.2, 1.5, 0.5 and 0.1 kb after *EcoRI* + *BamHI* digestion of  
5 the plasmid.

Although the present invention is described in the Examples listed above in terms of preferred  
embodiments, they are not to be regarded as limiting  
10 the scope of the invention. The above Examples serve to illustrate the principles and methodologies for constructing bacterial artificial chromosomes that can be introduced in a *Streptomyces* host. It will occur to those skilled in the art that selectable markers  
15 different from the *tsr* gene can be employed for selection in *Streptomyces*. Other useful markers are described in detail in laboratory manuals (Hopwood et al., 1985, Genetic Manipulation of *Streptomyces*: A Laboratory Manual, The John Innes Foundation, Norwich,  
20 UK) and include but are not limited to: genes conferring resistance to apramycin, kanamycin, erythromycin, hygromycin, viomycin. It will also occur to those skilled in the art that functions other than those specified by  $\Phi$ C31 can be used for directing  
25 site-specific integration in the *Streptomyces* chromosome. These functions are described in recent literature (Hopwood and Kieser, 1991, Methods Enzymol. 204:430- 458) and include but are not limited to those derived from pSAM2, SLP1, IS117. Bacterial artificial  
30 chromosomes derived from the *E. coli* F plasmid have been described (Shizuya et al., 1992, Proc. Natl. Acad. Sci. USA 89:8794-8797). It will occur to those skilled in the art that, using the principles and methodologies described above, the *int*-*tsr* cassette from pUIT4,  
35 prepared as described in Example 9, could be inserted into unique sites of pBAC108L (Shizuya et al., 1992,

BSI 0003 EPN

29

Proc. Natl. Acad. Sci. USA 89:8794-8797) or of suitable derivatives of this vector, leading to the formation of a BAC-based series of ESAC vectors. It will occur to those skilled in the art that other ESAC vectors  
5 differing, for example, in their size, in the *E. coli* replicon they carry, in the selectable marker for *E. coli*, in the cloning sites, can also be used in the present invention. Other differences and variations in the technical aspects of the present invention could be  
10 employed. These include but are not limited to: different methods and sources for obtaining selectable markers and integrative functions; different cloning sites and methodologies; different *E. coli* hosts for amplifying the recombinant constructs. All these  
15 variations fall within the scope of the present invention.

## 7.2 Construction of large inserts in ESAC

Two distinct methodologies for introducing large DNA  
20 fragments into the vectors described in Section 7.1 fall within the scope of the present invention. The first methodology can be referred to as the top-down approach and is depicted in Fig. 3. It consists of directly cloning the desired gene cluster into an ESAC  
25 vector through the construction of a genomic library of DNA fragments of average size of 100 kb, or more. The library is then screened with suitable probes (Section 7.3) in order to identify the desired cluster. The second methodology can be considered a bottom-up  
30 approach and is illustrated in Fig. 4. It consists of assembling the desired gene cluster from pre-existing smaller segments of cloned, overlapping DNA, through the iterative use of homologous recombination in *E. coli*. The desired overlapping clones encompass the  
35 desired gene cluster and are identified as described in Section 7.3. Both methodologies fall within the scope

BSI 0003 EPN

M 23.06.98

30

of this invention. Depending on a series of considerations, such as previous knowledge about the biosynthesis cluster, the extent of characterization of the producing strain, the existence of other natural products of interest produced by the original microorganism, one methodology may be preferred over the other. However, the two methodologies are not mutually exclusive and may also be pursued in parallel.

#### 7.2.1 Preparation of a large insert library

In order to prepare a large-insert library, particular care must be taken in the preparation of genomic DNA from the actinomycete strain of choice. Although several procedures have been described for the isolation of genomic DNA, very few are suitable for obtaining sufficient yields of high molecular weight DNA. The strain of choice is grown in a medium that allows dispersed growth to facilitate lysis of the cells. Examples of suitable growth media for different genera of actinomycetes can be found in the literature (Balows, Trueper, Dworkin, Harder and Schleifer eds., 1992, *The Prokaryotes*, 2nd edn., Springer-Verlag, New York, NY, USA). The growth time should be long enough for allowing the formation of a sufficient quantity of biomass; however, longer incubation times should be avoided, since mycelia are generally more resistant to lysis as they age. The mycelium is pelleted, washed and embedded in agarose for the subsequent lytic steps. Lysis of the cells is achieved by a combination of enzymatic (e.g., incubation with lysozyme and/or achromopeptidase) and mild physical treatments (e.g., SDS). The concentrations of reagents and the incubation times need to be optimized for each strain. A good starting point is represented by the conditions described in Example 12. The quality of the DNA preparation is checked by PFGE under appropriate

BSI 0003 EPN

31

conditions. Once a suitable preparation is obtained, the DNA can be digested as described in Example 13. The exact incubation time and the units of restriction endonuclease are adjusted to the particular DNA preparation for optimizing the size and yield of the bulk of digested DNA, which should exceed 150 kb. The partially digested DNA is size-fractionated on a PFGE gel, without exposure to ethidium bromide or UV light, in order to avoid damage to the DNA. The gel slice containing the desired DNA fraction is localized by staining the marker-containing portion of the gel and cut. All subsequent manipulations are performed with great care (Birren and Lai, 1993, Pulsed Field Gel Electrophoresis: A Practical Guide, Academic Press, New York, NY). The size-selected DNA is ligated to an appropriately prepared ESAC vector (see Example 14) employing a high molar excess of vector to insert (ca. 10:1) in order to minimize the formation of chimeric clones (i.e. those constituted by the religation of two unctiguous inserts). Subsequent steps are performed using published procedures for the cloning in bacterial artificial chromosomes, as described in Examples 16 and 17.

The genome size of actinomycetes is around 8 Mb. Consequently, a 10-genome equivalents library consisting of 800-clones with an average insert size of 100 kb has >99.9% probability of containing the desired clone ( (Sambrook et al., 1989, In *Molecular Cloning: A laboratory Manual*, 2nd edn, Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press). Therefore, the average clone in the library will have a 10-kb segment ( $8,000 \text{ kb} \div 800 \text{ clones} = 10 \text{ kb/clone}$ ) of unique DNA , i.e. DNA not found in any other clone. Consequently, a 90 kb cluster will have a high chance of being exactly contained within one or two 100-kb clones in a 800-clone library. The number of

BSI 0003 EPN

32

clones to be screened and the average insert size to be looked for in the ESAC library depends on the expected size of the biosynthesis gene cluster. The larger the difference between the average insert size and the expected size of the gene cluster, the smaller the number of clones to screen in order to identify the entire gene cluster in a single clone. ESAC DNA is prepared from a representative number of clones (from 24 to 48) obtained after electroporation of a ligation mixture and analyzed for determining the frequency of insert-carrying clones and their average size. If necessary, all insert containing clones can be analyzed by miniprep procedure (Birren and Lai, 1993, *Pulsed Field Gel Electrophoresis: A Practical Guide*, Academic Press, New York, NY, USA) and clones carrying inserts below a certain threshold can be discarded. Alternatively, the number of clones carrying insert of the appropriate size can be estimated after analysis of a representative number of ESAC clones. The quality of the library can be evaluated by probing with cloned genes from the strain (if available), or from highly conserved "housekeeping" genes from a strain with a similar GC content, such as *S. coelicolor*.

## Example 12

Preparation of high molecular weight chromosomal DNA  
*S. coelicolor* strain M145 is grown in YEME medium containing 0.5% (wt/vol) glycine for 40 h at 30°C on an orbital shaker (ca. 200 rpm). The mycelium is pelleted by centrifugation, washed with 10.3% sucrose and the chromosomal DNA is extracted from the mycelium embedded in 0.75% LMP agarose by treatment with 1 mg/ml lysozyme and with 1 mg/ml proteinase K in 0.1% SDS for 40 h at 50°C.

35

BSI 0003 EPN

33

Example 13

Preparation of partially digested chromosomal DNA  
*S. coelicolor* M145 chromosomal DNA, prepared as described in Example 12 and embedded in LMP agarose plugs, is partially digested by limiting the magnesium concentration for 20 min with 4 U of *Sau*3AI. The resulting DNA fragments are resolved by PFGE and the size-selected genomic DNA fraction (larger than 100 kb) is recovered and released from the agarose gel by digestion with gelase.

Example 14

Preparation of pPAC-S1 for library construction  
The vector pPAC-S1, prepared as described in Example 11, is cut with *Sca*I and then treated with calf intestinal phosphatase. The recovered DNA is then digested with *Bam*HI and treated with an excess of calf intestinal phosphatase. The short *Sca*I-*Bam*HI linker fragments are removed by spin dialysis.

Example 15Construction of the ESAC library

Size selected genomic DNA, prepared as described in Example 13, is ligated to pPAC-S1, prepared as described in Example 14, employing 300 Molecular Biology Units of T4 DNA ligase in a 50 µl final volume and using a ca. 10:1 molar ratio of vector to insert. The resulting ligation mixture contains the desired ESAC library, consisting of fragments *S. coelicolor* DNA inserted into the pPAC-S1 vector.

Example 16Introduction of the library into *E. coli* K12 DH10B

The ligation mixture, prepared as described in Example 15, is drop-dialyzed against 0.5 X TE for 2 h using 0.025 mm type VS membranes (Millipore) and a few µl are

BSI 0003 EPN

34

used to electroporate 40  $\mu$ l of electrocompetent *E. coli* DH10B cells. The electroporation conditions are: 2.5 kV, 100  $\Omega$  and 25 mFa employing the Biorad Gene Pulser II. The cells are plated on LB-agar plates containing

5 25  $\mu$ g/ml Km and 5% sucrose to select for recombinant cells harboring insert-carrying pPAC-S1. Individual colonies are picked into 0.1 ml of LB broth containing 25  $\mu$ g/ml Km in 96-well microtiter plates, where they are stored at -80 °C after overnight incubation and

10 addition of glycerol to 20% (v/v).

#### Example 17

##### Preparation of recombinant ESAC clones

Individual colonies, prepared as described in Example

15 16, are inoculated into 5 ml of LB broth containing 25  $\mu$ g/ml Km and grown overnight. ESAC DNA is isolated using the alkaline extraction procedure (Sambrook et al., 1989, In *Molecular Cloning: A laboratory Manual*, 2nd edn, Cold Spring Harbor, New York: Cold Spring

20 Harbor Laboratory Press)) without the phenol extraction step. The DNA is analyzed, after digestion with *Dra*I, by PFGE. Three bands of 7.4, 4.2 and 0.6 kb are common to all clones and represent vector DNA. The insert size in the recombinant ESAC clones is calculated by summing

25 up the sizes of the additional *Dra*I fragments and subtracting from this number 7.5 kb, the amount of the pPAC-S1 vector not included with the three *Dra*I fragments.

30 The examples described above illustrate the principles and methodologies of constructing a large-insert library of *S. coelicolor* DNA in an ESAC vector. Although the present invention is described in the Examples listed above in terms of preferred

35 embodiments, they are not to be regarded as limiting

BSI 0003 EPN

35

the scope of the invention. The above descriptions serve to illustrate the principles and methodologies for constructing a large-insert DNA library in an ESAC vector. It will occur to those skilled in the art that other *Streptomyces* strains can be used as a source of DNA for constructing the library. For example, an ESAC library of the rapamycin producer *Streptomyces hygroscopicus* ATCC 29253 could be constructed, employing the procedure reported for analyzing its DNA by PFGE (Ruan et al., 1997, Gene 203:1-9) and applying the principles and methodologies described in Examples 12 through 17.

It will also occur to those skilled in the art that actinomycete strains other than streptomycetes can be used as a source of DNA for constructing an ESAC library. These strains can belong to any genus of the order *Actinomycetales*, which include but are not limited to the genera reported in Table 1. Those skilled in the art understand that bacterial taxonomy is a rapidly evolving field and new genera may be described while old genera may be reclassified. Therefore, the list of bacteria genera related to actinomycetes is very likely to change. Nonetheless, the principles and methodologies of the present invention can be applied to any donor organism related to the actinomycetes.

It will also occur to those skilled in the art that different actinomycete strains will require growth media different from those reported in Example 12. Furthermore, it will occur to those skilled in the art that alternative media and conditions for growth can be employed for obtaining mycelia for DNA preparation; that alternative methods of lysis of mycelia can be utilized; that restriction endonucleases other than *Sau3AI* can be equally effective for constructing a library; that other methods for fragmenting DNA can also

BSI 0003 EPN

M 23.06.98

36

be employed. In addition, it will occur to those skilled in the art that ESAC vectors other than pPAC-S1, which include but are not limited to the possible vectors described in Section 7.1, can be used for

5 constructing a library. It will also occur to those skilled in the art that alternative methods for ligating DNA, for introducing the library in *E. coli* cells, and hosts other than DH10B are well described in the literature and can be employed in the present

10 invention. All the above variations in strains, reagents and methodologies that can be employed for preparing a large-insert library of actinomycete DNA into an ESAC vector fall within the scope of the present invention.

Table 1

List of exemplary genera of Actinomycetales

(http://www.ncbi.nlm.nih.gov/htbin-post/Taxonomy).

5	<i>Acidothe</i>	<i>planes</i>	<i>Nocardio</i>
	<i>rmus</i>	<i>Cellulomonas</i>	<i>oerskovia</i>
	<i>Actinobi</i>	<i>Chainia</i>	<i>Pelczaria</i>
	<i>spora</i>	<i>Clavibacter</i>	<i>Phenylobacteri</i>
10	<i>Actinoco</i>	65 <i>Coriobacterium</i>	120 <i>um</i>
	<i>rallia</i>	<i>Corynebacterium</i>	<i>Pilimelia</i>
	<i>Actinoki</i>	<i>Couchioplanes</i>	<i>Pimelobacter</i>
	<i>neospora</i>	<i>Cryobacterium</i>	<i>Planobispora</i>
	<i>Actinoma</i>	<i>Curtobacterium</i>	<i>Planomonospora</i>
15	<i>dura</i>	70 <i>Dactylosporangi</i>	125 <i>um</i>
	<i>Actinomy</i>	<i>m</i>	<i>Planopolyspora</i>
	<i>ces</i>	<i>Demetria</i>	<i>Planotetrasp</i>
	<i>Actinopl</i>	<i>Dermabacter</i>	<i>a</i>
	<i>anes</i>	<i>Dermacoccus</i>	<i>Prauseria</i>
20	<i>Actinopo</i>	75 <i>Dermatophilus</i>	130 <i>Promicromonosp</i>
	<i>lyspora</i>	<i>Dietzia</i>	<i>ora</i>
	<i>Actinopy</i>	<i>Elytrosporangium</i>	<i>Propionibacter</i>
	<i>cnidium</i>	<i>Excellospora</i>	<i>ium</i>
	<i>Actinosp</i>	<i>Exiguobacterium</i>	<i>Propioniferax</i>
25	<i>orangium</i>	80 <i>Frankia</i>	135 <i>Pseudonocardia</i>
	<i>Actinosy</i>	<i>Friedmanniella</i>	<i>Rarobacter</i>
	<i>nnema</i>	<i>Gardnerella</i>	<i>Rathayibacter</i>
	<i>Aeromicro</i>	<i>Geodermatophilus</i>	<i>Renibacterium</i>
	<i>obium</i>	<i>Glycomyces</i>	<i>Rhodococcus</i>
30	<i>Agrococc</i>	85 <i>Gordona</i>	140 <i>Rothia</i>
	<i>us</i>	<i>Herbidospora</i>	<i>Rubrobacter</i>
	<i>Agromyce</i>	<i>Intrasporangium</i>	<i>Saccharomonosp</i>
	<i>s</i>	<i>Janibacter</i>	<i>ora</i>
	<i>Ampullar</i>	<i>Jonesia</i>	<i>Saccharopolysp</i>
35	<i>iella</i>	90 <i>Kibdelosporangi</i>	145 <i>ora</i>
	<i>Amycolat</i>	<i>m</i>	<i>Saccharothrix</i>
	<i>a</i>	<i>Kineococcus</i>	<i>Sanguibacter</i>
	<i>Amycolat</i>	<i>Kineosporia</i>	<i>Skermania</i>
	<i>opsis</i>	<i>Kitasatoa</i>	<i>Spirilliplanes</i>
40	<i>Arcanoba</i>	95 <i>Kitasatosporia</i>	150 <i>Spirillospora</i>
	<i>cterium</i>	<i>Kocuria</i>	<i>Sporichthya</i>
	<i>Arthroba</i>	<i>Kutzneria</i>	<i>Stomatococcus</i>
	<i>cter</i>	<i>Kytococcus</i>	<i>Streptoallotei</i>
	<i>Atopobiu</i>	<i>Lentzea</i>	<i>chus</i>
45	<i>m</i>	100 <i>Luteococcus</i>	155 <i>Streptomyces</i>
	<i>Aureobac</i>	<i>Microbacterium</i>	<i>Streptosporang</i>
	<i>terium</i>	<i>Microbispora</i>	<i>ium</i>
	<i>Bifidoba</i>	<i>Micrococcus</i>	<i>Streptovertici</i>
	<i>cterium</i>	<i>Microellobospori</i>	<i>llium</i>
50	<i>Blastoco</i>	105 <i>a</i>	160 <i>Terrabacter</i>
	<i>ccus</i>	<i>Microlunatus</i>	<i>Terracoccus</i>
	<i>Bogoriel</i>	<i>Micromonospora</i>	<i>Thermoactinomy</i>
	<i>la</i>	<i>Microsphaera</i>	<i>ces</i>
	<i>Brachyba</i>	<i>Microtetraspora</i>	<i>Thermocrispum</i>
55	<i>cterium</i>	110 <i>Microthrix</i>	165 <i>Thermomonospor</i>
	<i>Brevibac</i>	<i>Mobiluncus</i>	<i>a</i>
	<i>terium</i>	<i>Mycobacterium</i>	<i>Tropheryma</i>
	<i>Catellat</i>	<i>Nesterenkonia</i>	<i>Tsukamurella</i>
	<i>ospora</i>	<i>Nocardia</i>	<i>Turicella</i>
170	60 <i>Catenulo</i>	115 <i>Nocardio</i>	
		<i>ides</i>	

BSI 0003 EPN

N 23.06.98

38

### 7.2.2 Assemblage by homologous recombination

The bottom-up strategy of assembling large fragment from pre-existing smaller segments of cloned DNA is described in this section. This methodology makes use of the same ESAC vectors described in the present invention under Section 7.1. The desired cluster is assembled from existing overlapping clones by the iterative use of homologous recombination in *E. coli*. In the example of Fig. 5, three overlapping clones, designated 1, 2 and 3, and derived from the genome of a donor organism, encompass the desired biosynthesis cluster. These clones include a leftward fragment "A" unique to clone 1; a fragment "B" common to clones 1 and 2; a fragment "C" common to clones 2 and 3; and a rightward fragment "D" unique to clone 3. The number of overlapping clones encompassing the cluster may vary. However, if  $n$  is the number of overlapping clones that cover the desired genomic segment, the number of fragments to consider will always be equal to  $n + 1$ . In the example illustrated in Fig. 5, the four fragments A, B, C and D are required. These fragments can range from a few hundred bp to a few kb, and are thus amenable to routine in vitro DNA manipulations. The cluster of Fig. 5 is reconstructed into an ESAC vector through the use of successive rounds of homologous recombination in *E. coli*. Fragments A and B (see Fig. 5) are cloned in a *ts* vector, as shown in Fig. 6, which carries a selectable marker,  $Cm^R$  as exemplified in Fig. 6. The same is done with fragment pairs B-C and C-D (Fig. 6). The relative orientation of each fragment pair in the *ts* vector must be the same as in the gene cluster. The fragments in each pair may be separated by a selectable marker, designated  $Ab^R$  in Fig. 6, to monitor interplasmid insert exchange. Therefore, three constructs in the *ts* vector, designated pAB1, pAB2 and

M 23.06.98

pAB3, are required. The A-B-C-D four-fragment cassette is cloned in an ESAC vector (Fig. 6). The relative orientation of the four fragments in the ESAC vector must be the same as in the gene cluster. Again, a

5 selectable marker may separate any of two fragments to monitor interplasmid insert exchange. The original clone (for example, a cosmid, which carries a selectable marker,  $Km^R$  as exemplified in Fig. 6) containing part of the cluster (Fig. 5) and the cognate

10 *ts* construct (Fig. 6) are introduced into the same *E. coli* cell. The interplasmid cointegrate between the original clone and the *ts* construct is selected at the non-permissive temperature for the *ts* replicon. This occurs through a single, reciprocal homologous

15 recombination mediated by either one of the two fragments in the A-B, B-C or C-D pairs. The cointegrate is then resolved at the permissive temperature, leading to insert exchange between the two replicons (see Fig. 7). The presence in the *ts* replicon of the genomic

20 segment comprised between fragments A and B can be monitored by the appearance of  $Cm^R Ab^S$  colonies. This is done for clone 1 and pAB1, resulting in pAB2; for clone 2 and pBC1, resulting in pBC2; and for clone 3 and pCD1, resulting in pCD2. Each insert from the

25 original overlapping clones (Fig. 5) is thus transferred into the *ts* replicon, as outlined in Fig. 7. Subsequently, the inserts from clone 1, now present in the *ts* plasmid pAB2, is introduced into the ESAC construct carrying the entire A-B-C-D cassette. This is

30 done by selecting for the interplasmid cointegrate between the pAB2 and the ESAC construct at the non-permissive temperature, and then resolving the cointegrate at the permissive temperature, selecting for  $Km^R Ab^S$  colonies. This leads to insert exchange

35 between the two replicons (as shown in Fig. 8). Next, a selectable marker is introduced in the growing ESAC

BSI 0003 EPN

M 23.06.98

40

clone between the next fragment pair, again through the use of two rounds of single, reciprocal homologous recombination mediated by plasmid pBC1, leading to the appearance of  $Km^R$   $Ab^R$  colonies. Subsequently, the interplasmid exchange with pBC2 leads to the introduction of the genomic segment comprised between fragments B and C. Finally, the use of pCD1 first and subsequently of pCD2 leads to the reconstruction of the genomic segment into the ESAC vector. Therefore, through the use of alternating steps, the  $Ab^R$  marker first and the genomic segment later are introduced between any fragment pair, as schematized in Fig. 8. This iterative procedure results in the reconstruction of the original chromosomal region in the ESAC vector.

A series of examples described herein illustrate how the 90-kb gene cluster involved in the biosynthesis of the antibiotic GE2270 is assembled from three pre-existing cosmids via homologous recombination. The cosmids, designated pRP16, pRP31 and pRP58, are identified in a cosmid library constructed in the vector Lorist6 by the use of selective hybridization probes. The relevant information about the cluster is reported in Fig. 9. The reconstruction of the cluster results in the formation of the intermediate ESAC derivatives pPAD1, pPAD2, pPAD4 and pPAD6, carrying inserts of 10, 39, 68 and 89 kb, respectively. The examples reported herein serve to illustrate the principles and methodologies of the present invention and are not meant to restrict its scope.

#### Example 18

##### Isolation of cosmid clones pRP16, pRP31 and pRP58

A cosmid library of *P. rosea* DNA prepared in the vector Lorist6 is screened with oligonucleotide probes Pep6 and Pep8, according to the conditions described under General Methods, Section 6. Among the positive colonies

BSI 0003 EPN

M 20.06.98

41

identified, several cosmids were found to span the ca. 90 kb segment of the *P. rosea* chromosome reported in Fig. 9. Signature sequences close to the left and right end of the segment of Fig. 9 are reported in Fig. 10A and 10B, respectively. Three cosmids are chosen for further studies. Cosmid pRP16 exhibits, after digestion with *Bam*HI and resolution by agarose gel-electrophoresis, fragments of 7.5, 7.2, 5.6, 5.2, 2.7, 2.0, 1.9, 1.9, 1.8, 1.6, 1.4, 0.9 and 0.7 kb. Cosmid pRP31 exhibits, after digestion with *Bam*HI and resolution by agarose gel-electrophoresis, fragments of 10.5, 6.2, 3.1, 2.8, 2.6, 2.5, 2.1, 1.9, 1.9, 1.5, 1.4, 1.2, 1.0, 1.0, 0.9, 0.9, 0.7, 0.6, 0.5, 0.5, 0.1 and 0.1 kb. Cosmid pRP58 exhibits, after digestion with *Bam*HI and resolution by agarose gel-electrophoresis, fragments of 10.0, 7.6, 6.7, 6.2, 3.4, 3.0, 2.8, 2.1, 1.0, 1.0, 0.9, 0.9, 0.7, 0.6, 0.5, 0.5, 0.3 and 0.1 kb.

#### Example 19

##### Construction of plasmid pUA1

The 0.9 kb *Sma*I-*Sst*I fragment, comprised between map coordinates 2.0-2.9 kb of Fig. 9, is obtained from cosmid pRP16, prepared as described in Example 18, and ligated to pUC18 previously digested with *Sst*I and *Sma*I. The resulting mixture contains the desired plasmid pUA1.

#### Example 20

##### Construction of *E. coli* K12 XL1blue/pUA1

Approximately 10 ng of plasmid pUA1, prepared as described in Example 19, are used to transform *E. coli* XL1blue and a few of the resulting Ap<sup>R</sup> colonies that appear on the LB-agar plates are analyzed for their plasmid content. One colony is found to carry pUA1, as verified by the observation, upon agarose gel-electrophoresis, of fragments of 2.7 and 0.9 kb after

BSI 0003 EPN

M 23.06.98

42

digestion of the plasmid with *Bam*HI + *Sst*I.

Example 21

Construction of plasmid pUA2

- 5 The 0.9 kb *Bam*HI-*Sst*I fragment from pUA1, prepared as described in Example 20, is ligated to pUCBM20 previously digested with *Bam*HI and *Sst*I. The resulting mixture contains the desired plasmid pUA2.

10

Example 22

Construction of *E. coli* K12 XL1blue/pUA2

- Approximately 10 ng of plasmid pUA2, prepared as described in Example 20, are used to transform *E. coli* XL1blue and a few of the resulting Ap<sup>R</sup> colonies that appear on the LB-agar plates are analyzed for their plasmid content. One colony is found to carry pUA2, as verified by the observation, upon agarose gel-electrophoresis, of fragments of 2.7 and 0.9 kb after digestion of the plasmid with *Eco*RI + *Sst*I.

20

Example 23

Construction of plasmid pUB1

- The 1.8 kb *Sst*I-*Bam*HI fragment, comprised between map coordinates 33.4-35.2 of Fig. 9, is obtained from cosmid pRP16, prepared as described in Example 18, and ligated to pUC18 previously digested with *Sst*I + *Bam*HI. The ligation mixture contains the desired plasmid pUB1.

30

Example 24

Construction of *E. coli* K12 XL1blue/pUB1

- Approximately 10 ng of plasmid pUB1, prepared as described in Example 23, are used to transform *E. coli* XL1blue and a few of the resulting Ap<sup>R</sup> colonies that appear on the LB-agar plates are analyzed for their plasmid content. One colony is found to contain plasmid pUB1 as verified by the observation, upon agarose gel

35

BSI 0003 EPN

11.23.05.99

43

electrophoresis, of fragments 2.7 and 1.8 kb after digestion with *Sst*I + *Xba*I.

#### Example 25

##### 5                   Construction of plasmid pUC1

The 6.2 kb *Bam*HI fragment, comprised between map coordinates 54.2-60.4 kb of Fig. 9, is obtained from cosmid pRP58, prepared as described in Example 18, and ligated to pUC18 previously digested with *Bam*HI. The  
10 ligation mixture contains the desired plasmid pUC1.

#### Example 26

##### Construction of *E. coli* K12 XL1blue/pUC1

Approximately 10 ng of plasmid pUC1, prepared as  
15 described in Example 25, are used to transform *E. coli* XL1blue and a few of the resulting Ap<sup>R</sup> colonies that appear on the LB-agar plates are analyzed for their plasmid content. One colony is found to carry pUC1, as verified by the observation, upon agarose gel-  
20 electrophoresis, of fragments of 4.9 and 4.0 kb after digestion of the plasmid with *Pst*I.

#### Example 27

##### Construction of plasmid pUD1

25 Synthetic oligonucleotides 5'-GATCTAAGCTTGGGGG-3' and 5'-CCCCCAAGCTTA-3' are annealed and ligated to the 1.5 kb *Pst*I-*Bam*HI fragment, comprised between map coordinates 89.5-91.0 kb of Fig. 9 and obtained from cosmid pRP58, prepared as described in Example 18. The  
30 ligation mixture is digested with *Hind*III and ligated to pUC18 previously digested with *Pst*I + *Hind*III. The resulting mixture contains the desired plasmid pUD1.

#### Example 28

##### 35                   Construction of *E. coli* K12 XL1blue/pUD1

Approximately 10 ng of plasmid pUD1, prepared as

BSI 0003 EPN

11 23 06 98

44

described in Example 27, are used to transform *E. coli* XL1blue and a few of the resulting Ap<sup>R</sup> colonies that appear on the LB-agar plates are analyzed for their plasmid content. One colony is found to contain plasmid  
5 pUD1 as verified by the observation, upon agarose gel-electrophoresis, of fragments of 2.7 and 1.5 kb after digestion with *Pst*I + *Hind*III.

#### Example 29

##### 10 Construction of plasmid pUAB1

The 0.9 kb *Eco*RI-*Sst*I fragment from plasmid pUA2, prepared as described in Example 22, and the 1.8 kb *Sst*I-*Bam*HI fragment from pUB1, prepared as described in  
15 Example 24, are ligated to pUC18 previously digested with *Eco*RI + *Bam*HI. The ligation mixture contains the desired plasmid pUAB1.

#### Example 30

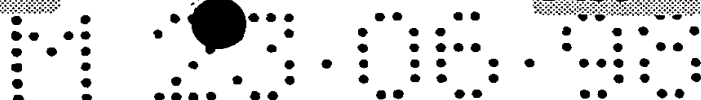
##### Construction of *E. coli* K12 XL1blue/pUAB1

20 Approximately 10 ng of plasmid pUAB1, prepared as described in Example 29, are used to transform *E. coli* XL1blue and a few of the resulting Ap<sup>R</sup> colonies that appear on the LB-agar plates are analyzed for their plasmid content. One colony is found to carry pUAB1, as  
25 verified by the observation, upon agarose gel-electrophoresis, of two fragments of 2.7 kb after digestion of the plasmid with *Eco*RI + *Xba*I.

#### Example 31

##### 30 Isolation of the *tetR* fragment

The 1.6 kb fragment containing the *tetR* gene is isolated after PCR amplification of pBR322 DNA using oligonucleotides 5'-GAGCTCTCATGTTTGACAGCT-3' and 5'-GAGCTCTGACTTCCGCGTTTCCAG-3' as primers, followed by  
35 digestion with *Sst*I.

Example 32Construction of plasmid pUAB2

Plasmid pUAB1, prepared as described in Example 30, is digested with *Sst*I and ligated to the *tet*R fragment prepared as described in Example 31. The ligation mixture contains the desired plasmid pUAB2.

Example 33Construction of *E. coli* K12 DH5 $\alpha$ /pUAB2

Approximately 10 ng of plasmid pUAB2, prepared as described in Example 32, are used to transform *E. coli* DH5 $\alpha$  and a few of the resulting Tc<sup>R</sup> Ap<sup>R</sup> colonies that appear on the LB-agar plates are analyzed for their plasmid content. One colony is found to carry pUAB2, as verified by the observation, upon agarose gel-electrophoresis, of fragments of 4.3 and 2.7 kb after digestion of the plasmid with *Eco*RI + *Xba*I.

Example 34Construction of plasmid pUBC1

The 1.8 kb *Sst*I-*Xba*I fragment obtained from plasmid pUB1, prepared as described in Example 24, and the 4.0 kb *Xba*I-*Pst*I fragment obtained from plasmid pUC1, prepared as described in Example 26, are ligated to pUC18 previously digested with *Sst*I + *Pst*I. The ligation mixture contains the desired plasmid pUBC1.

Example 35Construction of *E. coli* K12 XL1blue/pUBC1

Approximately 10 ng of plasmid pUBC1, prepared as described in Example 34, are used to transform *E. coli* XL1blue and a few of the resulting Ap<sup>R</sup> colonies that appear on the LB-agar plates are analyzed for their plasmid content. One colony is found to carry pUBC1, as verified by the observation, upon agarose gel-electrophoresis, of fragments of 5.8 and 2.7 kb after

BSI 0003 EPN

M 23.06.98

46

digestion of the plasmid with *EcoRI* + *HindIII*.

Example 36

Construction of plasmid pUBC2

5 Plasmid pUBC1, prepared as described in Example 35 and previously digested with *XbaI*, and the *tetR* fragment, prepared as described in Example 31, are treated with T4 DNA polymerase and T4 DNA ligase. The ligation mixture contains the desired plasmid pUBC2.

10

Example 37

Construction of *E. coli* K12 DH5 $\alpha$ /pUBC2

Approximately 10 ng of plasmid pUBC2, prepared as described in Example 36, are used to transform *E. coli* DH5 $\alpha$  and a few of the resulting Tc<sup>R</sup> Ap<sup>R</sup> colonies that appear on the LB-agar plates are analyzed for their plasmid content. One colony is found to carry pUBC2, as verified by the observation, upon agarose gel-electrophoresis, of fragments of 5.6 and 4.5 kb after digestion of the plasmid with *HindIII*.

20

Example 38

Construction of plasmid pUCD1

The 4.0 kb *XbaI*-*PstI* fragment obtained from plasmid pUC1, prepared as described in Example 26, and the 1.5 kb *PstI*-*HindIII* fragment isolated from plasmid pUD1, prepared as described in Example 28, are ligated to pUC18 previously digested with *XbaI* and *HindIII*. The mixture contains the desired plasmid pUCD1.

30

Example 39

Construction of *E. coli* K12 XL1blue/pUCD1

Approximately 10 ng of plasmid pUCD1, prepared as described in Example 38, are used to transform *E. coli* XL1blue and a few of the resulting Ap<sup>R</sup> colonies that appear on the LB-agar plates are analyzed for their

35

BSI 0003 EPN

47

plasmid content. One colony is found to carry pUCD1, as verified by the observation, upon agarose gel-electrophoresis, of fragments of 5.5 and 2.7 kb after digestion of the plasmid with *Xba*I + *Hind*III.

5

#### Example 40

##### Construction of plasmid pUCD2

Plasmid pUCD1, prepared as described in Example 39 and previously digested with *Pst*I, and the *tet*R fragment prepared as described in Example 31, are treated with T4 DNA polymerase and T4 DNA ligase. The ligation mixture contains the desired plasmid pUCD2.

10

#### Example 41

##### Construction of *E. coli* K12 DH5 $\alpha$ /pUCD2

Approximately 10 ng of plasmid pUCD2, prepared as described in Example 40, are used to transform *E. coli* DH5 $\alpha$  and a few of the resulting Tc<sup>R</sup> Ap<sup>R</sup> colonies that appear on the LB-agar plates are analyzed for their plasmid content. One colony is found to carry pUCD2, as verified by the observation, upon agarose gel-electrophoresis, of fragments of 6.7 and 3.1 kb after digestion of the plasmid with *Hind*III.

20

25

#### Example 42

##### Construction of plasmid pUAD1

The 4.3 kb *Eco*RI-*Xba*I fragment obtained from plasmid pUAB2, prepared as described in Example 33, and the 5.5 *Xba*I-*Hind*III fragment from plasmid pUCD1, prepared as described in Example 39, are ligated to pUC18, previously digested with *Eco*RI + *Hind*III. The ligation mixture contains the desired plasmid pUAD1.

30

#### Example 43

##### Construction of *E. coli* K12 DH5 $\alpha$ /pUAD1

Approximately 10 ng of plasmid pUAD1, prepared as

35

BSI 0003 EPN

48

described in Example 42, are used to transform *E. coli* DH5 $\alpha$  and a few of the resulting Tc<sup>R</sup> Ap<sup>R</sup> colonies that appear on the LB-agar plates are analyzed for their plasmid content. One colony is found to carry pUAD1, as  
5 verified by the observation, upon agarose gel-electrophoresis, of fragments of 8.9 and 3.6 kb after digestion of the plasmid with *Hind*III.

#### Example 44

##### 10 Construction of plasmid pMAB1

The 4.3 kb *Eco*RI-*Xba*I fragment obtained from plasmid pUAB2, prepared as described in Example 33, is treated with T4 DNA Polymerase and ligated to pMAK705 previously digested with *Hinc*II. The ligation mixture  
15 contains the desired plasmid pMAB1.

#### Example 45

##### Construction of *E. coli* K12 C600/pMAB1

Approximately 10 ng of plasmid pMAB1, prepared as  
20 described in Example 43, are used to transform *E. coli* C600 and a few of the resulting Cm<sup>R</sup> Tc<sup>R</sup> colonies that appear on the LB-agar plates are analyzed for their plasmid content. One colony is found to carry pMAB1, as verified by the observation, upon agarose gel-  
25 electrophoresis, of fragments of 4.1, 3.4, 1.4 and 0.9 kb after digestion of the plasmid with *Hind*III + *Eco*RI.

#### Example 46

##### Construction of plasmid pMBC1

30 The 7.1 kb fragment from plasmid pUBC2, prepared as described in Example 37, is obtained after partial digestion with *Pst*I, treated with T4 DNA polymerase and ligated to pMAK705 previously digested with *Hinc*II. The ligation mixture contains the desired plasmid pMBC1.

35

BSI 0003 EPN

M 23.06.98

49

Example 47Construction of *E. coli* K12 C600/pMBC1

Approximately 10 ng of plasmid pMBC1, prepared as described in Example 46, are used to transform *E. coli* C600 and a few of the resulting  $\text{Cm}^R \text{Tc}^R$  colonies that appear on the LB-agar plates are analyzed for their plasmid content. One colony is found to carry pMBC1, as verified by the observation, upon agarose gel-electrophoresis, of fragments of 9.5, 1.5, 1.3 and 0.3 kb after digestion of the plasmid with *Bam*HI.

Example 48Construction of plasmid pMCD1

The 7.1 kb fragment from plasmid pUCD2, prepared as described in Example 41, is obtained by complete digestion with *Xba*I and partial digestion with *Hind*III, treated with T4 DNA polymerase and ligated to pMAK705, previously digested with *Hinc*II. The ligation mixture contains the desired plasmid pMCD1.

Example 49Construction of *E. coli* K12 C600/pMCD1

Approximately 10 ng of plasmid pMCD1, prepared as described in Example 48, are used to transform *E. coli* C600 and a few of the resulting  $\text{Cm}^R \text{Tc}^R$  colonies that appear on the LB-agar plates are analyzed for their plasmid content. One colony is found to carry pMCD1, as verified by the observation, upon agarose gel-electrophoresis, of fragments of 8.6 and 4.3 kb after digestion of the plasmid with *Bam*HI.

Example 50Construction of plasmid pPAD1

The 10.0 kb *Eco*RI-*Nde*I fragment from plasmid pUAD1, prepared as described in Example 43, is ligated to pPAC-S1, prepared as described in Example 11 and

BSI 0003 EPN

M 23.06.98

50

previously digested with *ScaI*. The ligation mixture contains the desired plasmid pPAD1.

#### Example 51

##### 5        Construction of *E. coli* K12 C600/pPAD1

Approximately 10 ng of plasmid pPAD1, prepared as described in Example 50, are used to transform *E. coli* C600 and a few of the resulting Km<sup>R</sup> Tc<sup>R</sup> colonies that appear on the LB-agar plates are analyzed for their  
10        plasmid content. One colony is found to carry pPAD1, as verified by the observation, upon agarose gel-electrophoresis, of fragments of 19.7, 5.8, 3.1 and 1.2 kb after digestion of the plasmid with *Bam*HI. After digestion with *DraI* and resolution by PFGE, pPAD1  
15        yields fragments of 17.4, 7.4, 4.2 and 0.6 kb.

#### Example 52

##### Construction of *E. coli* K12 C600/pMAB1::pRP16

*E. coli* C600/pMAB1, prepared as described in Example  
20        45, is transformed with ca. 50 ng of pRP16, prepared as described in Example 18. The Cm<sup>R</sup> Km<sup>R</sup> colonies that appear at 30 °C on the LB-agar plates are grown at 30°C in LB broth containing Km and Cm, aliquots are withdrawn at various times and appropriate dilutions  
25        plated. Few of the Cm<sup>R</sup> Km<sup>R</sup> colonies that appear on the LB-agar plates after overnight incubation at 44°C are grown in LB broth containing Km and Cm for 16 h at 44°C and analyzed for their plasmid content. One colony is found to carry pMAB1::pRP16, as verified by the  
30        observation, upon agarose gel-electrophoresis, of fragments of 34, 10.7, 1.7, 1.6, 1.5, 1.2 and 0.6 kb after digestion of the plasmid with *Eco*RI.

#### Example 53

##### 35        Construction of *E. coli* K12 C600/pMAB2

Several colonies of *E. coli* C600/pMAB1::pRP16, prepared

BSI 0003 EPN

51

as described in Example 52, are grown individually in LB broth containing Cm for 24 h at 30°C, diluted 1:100 and incubated for further 8 h. Appropriate dilutions are plated. Few of the resulting Cm<sup>R</sup> Km<sup>S</sup> Tc<sup>S</sup> colonies that appear at 30°C are analyzed for their plasmid content. One colony is found to carry pMAB2, as verified by the observation, upon agarose gel-electrophoresis, of fragments of 37 and 1.5 kb after digestion of the plasmid with *EcoRI*.

#### Example 54

##### Construction of *E. coli* K12 DH1/pMBC1::pRP31

Approximately 50 ng of pRP31, prepared as described in Example 18, are used to transform *E. coli* DH1 cells. Competent cells are prepared from one of the resulting Km<sup>R</sup> colonies and transformed with ca. 10 ng of plasmid pMBC1, prepared as described in Example 47. The Cm<sup>R</sup> Km<sup>R</sup> colonies that appear at 30°C on the LB-agar plates are grown at 30°C in LB broth containing Km and Cm, aliquots are withdrawn at various times and appropriate dilutions plated. Few of the Cm<sup>R</sup> Km<sup>R</sup> colonies that appear on the LB-agar plates after overnight incubation at 44°C are grown in LB broth containing Km and Cm for 16 h at 44°C and analyzed for their plasmid content. One colony is found to carry pMBC1::pRP31, as verified by the observation, upon agarose gel-electrophoresis, of fragments of 22.2, 14.1, 14.0 and 6.0 kb after digestion of the plasmid with *EcoRV*.

#### Example 55

##### Construction of *E. coli* K12 DH1/pMBC2

Several colonies of *E. coli* DH1/pMBC1::pRP31, prepared as described in Example 54, are grown individually in LB broth containing Cm for 24 h at 30°C, diluted 1:100 and incubated for further 8 h. Appropriate dilutions

BSI 0003 EPN

M 23.06.98

52

are plated. Few of the resulting  $\text{Cm}^R \text{Km}^S \text{Tc}^S$  colonies that appear at 30°C are analyzed for their plasmid content. One colony is found to carry pMBC2, as verified by the observation, upon agarose gel-electrophoresis, of fragments of 14.4, 14.1 and 1.5 kb after digestion of the plasmid with *EcoRI*.

#### Example 56

##### Construction of *E. coli* K12 DH1/pMCD1::pRP58

10 Approximately 50 ng of pRP58, prepared as described in Example 18, are used to transform *E. coli* DH1 cells. Competent cells are prepared from one of the resulting  $\text{Km}^R$  colonies and transformed with ca. 10 ng of plasmid pMCD1, prepared as described in Example 48. The  $\text{Cm}^R \text{Km}^R$  colonies that appear at 30°C on the LB-agar plates are grown at 30°C in LB broth containing Km and Cm, aliquots are withdrawn at various times and appropriate dilutions plated. Few of the  $\text{Cm}^R \text{Km}^R$  colonies that appear on the LB-agar plates after overnight incubation at 44°C are grown in LB broth containing Km and Cm for 16 h at 44°C and analyzed for their plasmid content. One colony is found to carry pMCD1::pRP58, as verified by the observation, upon agarose gel-electrophoresis, of fragments of 39, 16, 1.7, 1.6, 1.5, 1.2 and 0.6 kb after digestion of the plasmid with *EcoRI*.

#### Example 57

##### Construction of *E. coli* K12 DH1/pMCD2

Several colonies of *E. coli* DH1/pMCD1::pRP58, prepared as described in Example 56, are grown individually in LB broth containing Cm for 24 h at 30°C, diluted 1:100 and incubated for further 8 h. Appropriate dilutions are plated. Few of the resulting  $\text{Cm}^R \text{Km}^S \text{Tc}^S$  colonies that appear at 30°C are analyzed for their plasmid content. One colony is found to carry pMCD2, as

BSI 0003 EPN

M 23.06.98

53

verified by the observation, upon agarose gel-electrophoresis, of fragments of 42 and 1.5 kb after digestion of the plasmid with *EcoRI*.

5

Example 58Construction of *E. coli* K12 C600/pMAB2::pPAD1

*E. coli* C600/pMAB2, prepared as described in Example 53, is transformed with ca. 50 ng of plasmid pPAD1, prepared as described in Example 51. The  $\text{Cm}^R$   $\text{Km}^R$  colonies that appear at 30°C on the LB-agar plates are grown at 30°C in LB broth containing Km and Cm, aliquots are withdrawn at various times and appropriate dilutions plated. Few of the  $\text{Cm}^R$   $\text{Km}^R$  colonies that appear on the LB-agar plates after overnight incubation at 44°C are grown in LB broth containing Km and Cm for 16 h at 44°C and analyzed for their plasmid content. One colony is found to carry pMAB2::pPAD1, as verified by the observation, upon agarose gel-electrophoresis, of fragments of 19.7, 7.2, 5.6, 5.6, 5.5, 5.2, 3.1, 2.7, 1.9, 1.9, 1.8, 1.8, 1.6, 1.4, 1.2, 0.9, 0.9 and 0.7 kb after digestion of the plasmid with *BamHI*.

Example 59Construction of *E. coli* K12 C600/pPAD2

Several colonies of *E. coli* C600/pMAB2::pPAD1, prepared as described in Example 58, are grown individually in LB containing Km for 24 h at 30°C, diluted 1:100 and incubated for further 8 h at 44°C. Appropriate dilutions are plated. Few of the resulting  $\text{Km}^R$   $\text{Cm}^S$   $\text{Tc}^S$  colonies that appear at 37°C are analyzed for their plasmid content. One colony is found to carry pPAD2, as verified by the observation, upon agarose gel-electrophoresis, of fragments of 19.7, 7.2, 5.6, 5.5, 5.2, 2.7, 1.9, 1.9, 1.8, 1.8, 1.6, 1.4, 0.9 and 0.7 kb after digestion of the plasmid with *BamHI*. After *DraI*

BSI 0003 EPN

M 23.08.98

54

digestion and resolution by PFGE, pPAD2 yields fragments of 45, 7.4, 4.2 and 0.6 kb.

#### Example 60

5

##### Construction of plasmid pMCD3

The 1.4 kb *KpnI*-*XhoII* fragment obtained from plasmid pCYPAC2 after digestion with *XhoII*, treatment with T4 DNA polymerase and digestion with *KpnI*, and the 7.1 kb *XbaI*-*HindIII* fragment from pUCD2, prepared as described in Example 40 and obtained after partial digestion with *HindIII*, complete digestion with *XbaI* and treatment with T4 DNA polymerase, are ligated to pMAK705, previously digested with *KpnI* + *HincII*. The ligation mixture contains the desired plasmid pMCD3.

15

#### Example 61

##### Construction of *E. coli* K12 C600/pMCD3

Approximately 10 ng of plasmid pMCD3, prepared as described in Example 60, are used to transform *E. coli* C600 and a few of the resulting  $\text{Cm}^R$   $\text{Tc}^R$  colonies that appear on the LB-agar plates are analyzed for their plasmid content. One colony is found to carry pMCD3, as verified by the observation, upon agarose gel-electrophoresis, of fragments of 9.8 and 4.3 kb after digestion of the plasmid with *BamHI*.

25

#### Example 62

##### Construction of *E. coli* K12 C600/pPAD2::pMCD3

*E. coli* C600/pPAD2, prepared as described in Example 59, is transformed with ca. 10 ng of plasmid pMCD3, prepared as described in Example 61. The  $\text{Cm}^R$   $\text{Km}^R$  colonies that appear at 30°C on the LB-agar plates are grown at 30°C in LB broth containing Km and Cm, aliquots are withdrawn at various times and appropriate dilutions plated. Few of the  $\text{Cm}^R$   $\text{Km}^R$  colonies that appear on the LB-agar plates after overnight incubation

35

BSI 0003 EPN

11.23.06.98

55

at 44°C are grown in LB broth containing Km and Cm for 16 h at 44°C and analyzed for their plasmid content. One colony is found to carry pPAD2::pMCD3, as verified by the observation, upon agarose gel-electrophoresis, of fragments of 19.7, 9.8, 7.2, 5.6, 5.5, 5.2, 4.3, 2.7, 1.9, 1.9, 1.8, 1.8, 1.6, 1.4, 0.9 and 0.7 kb after digestion of the plasmid with *Bam*HI.

#### Example 63

##### 10      Construction of *E. coli* K12 C600/pPAD3

Several colonies of *E. coli* C600/pPAD2::pMCD3, prepared as described in Example 62, are grown individually in LB broth containing Km for 24 h at 30°C, diluted 1:100 and incubated for further 8 h at 44°C. Appropriate dilutions are plated. Few of the resulting Km<sup>R</sup> Cm<sup>S</sup> Tc<sup>R</sup> colonies are analyzed for their plasmid content. One colony is found to carry pPAD3, as verified by the observation, upon agarose gel-electrophoresis, of fragments of 21.5, 7.2, 5.6, 5.2, 4.3, 2.7, 1.9, 1.9, 1.8, 1.8, 1.6, 1.4, 0.9 and 0.7 kb after digestion of the plasmid with *Bam*HI.

#### Example 64

##### Construction of *E. coli* K12 C600/pPAD3::pMCD2

25 *E. coli* C600/pPAD3, prepared as described in Example 63, is transformed with ca. 50 ng of plasmid pMCD2, prepared as described in Example 57. The Cm<sup>R</sup> Km<sup>R</sup> colonies that appear at 30°C on the LB-agar plates are grown at 30°C in LB broth containing Km and Cm, aliquots are withdrawn at various times and appropriate dilutions plated. Few of the Cm<sup>R</sup> Km<sup>R</sup> colonies that appear on the LB-agar plates after overnight incubation at 44°C are grown in LB broth containing Km and Cm for 16 h at 44°C and analyzed for their plasmid content.

BSI 0003 EPN

M 23.06.98

56

One colony is found to carry pPAD3::pMCD2, as verified by the observation, upon agarose gel-electrophoresis, of fragments of 21.5, 10, 9.0, 7.6, 7.2, 6.2, 5.6, 5.2, 4.3, 3.1, 2.8, 2.7, 1.9, 1.9, 1.8, 1.8, 1.6, 1.4, 1.0, 0.9, 0.9, 0.9, 0.7, 0.5, 0.3 and 0.1 kb after digestion of the plasmid with BamHI.

#### Example 65

##### Construction of *E. coli* K12 C600/pPAD4

Several colonies of *E. coli* C600/pPAD3::pMCD2, prepared as described in Example 64, are grown individually in LB broth containing Km for 24 h at 30°C, diluted 1:100 and incubated for further 8 h at 44°C. Appropriate dilutions are plated. Few of the resulting Km<sup>R</sup> Cm<sup>S</sup> Tc<sup>S</sup> colonies are analyzed for their plasmid content. One colony is found to carry pPAD4, as verified by the observation, upon agarose gel-electrophoresis, of fragments of 22, 10, 7.6, 7.2, 6.2, 5.6, 5.2, 3.1, 2.8, 2.7, 1.9, 1.9, 1.8, 1.8, 1.6, 1.4, 1.0, 0.9, 0.9, 0.9, 0.7, 0.5, 0.3 and 0.1 kb after digestion of the plasmid with BamHI. After DraI digestion and resolution by PFGE, pPAD4 yields fragments of 79, 4.2 and 0.6 kb.

#### Example 66

##### Construction of *E. coli* K12 C600/pPAD4::pMBC1

*E. coli* C600/pPAD4, prepared as described in Example 65, is transformed with ca. 10 ng of plasmid pMBC1, prepared as described in Example 47. The Cm<sup>R</sup> Km<sup>R</sup> colonies that appear at 30°C on the LB-agar plates are grown at 30°C in LB broth containing Km and Cm, aliquots are withdrawn at various times and appropriate dilutions plated. Few of the Cm<sup>R</sup> Km<sup>R</sup> colonies that appear on the LB-agar plates after overnight incubation at 44°C are grown in LB broth containing Km and Cm for 16 h at 44°C and analyzed for their plasmid content.

BSI 0003 EPN

57

One colony is found to carry pPAD4::pMBC1, as verified by the observation, upon agarose gel-electrophoresis, of fragments of 22, 10, 9.6, 7.6, 7.2, 6.2, 5.6, 5.2, 3.1, 2.8, 2.7, 1.9, 1.9, 1.8, 1.8, 1.6, 1.5, 1.4, 1.3, 1.0, 0.9, 0.9, 0.9, 0.7, 0.5, 0.3, 0.3 and 0.1 kb after digestion of the plasmid with *Bam*HI.

#### Example 67

##### Construction of *E. coli* K12 C600/pPAD5

10 Several colonies of *E. coli* C600/pPAD4::pMBC1, prepared as described in Example 66, are grown individually in LB broth containing Km for 24 h at 30°C, diluted 1:100 and incubated for further 8 h at 44°C. Appropriate dilutions are plated. Few of the resulting Km<sup>R</sup> Cm<sup>S</sup> Tc<sup>R</sup> colonies are analyzed for their plasmid content. One colony is found to carry pPAD5, as verified by the observation, upon agarose gel-electrophoresis, of fragments of 22, 10, 7.6, 7.2, 6.2, 5.6, 5.2, 3.1, 2.8, 2.7, 1.9, 1.9, 1.8, 1.8, 1.6, 1.4, 1.3, 1.0, 0.9, 0.9, 0.9, 0.7, 0.5, 0.3, 0.3 and 0.1 kb after digestion of the plasmid with *Bam*HI.

#### Example 68

##### Construction of *E. coli* K12 C600/pPAD5::pMBC2

25 *E. coli* C600/pPAD5, prepared as described in Example 67, is transformed with ca. 50 ng of plasmid pMBC2, prepared as described in Example 55. The Cm<sup>R</sup> Km<sup>R</sup> colonies that appear at 30°C on the LB-agar plates are grown at 30°C in LB broth containing Km and Cm, aliquots are withdrawn at various times and appropriate dilutions plated. Few of the Cm<sup>R</sup> Km<sup>R</sup> colonies that appear on the LB-agar plates after overnight incubation at 44°C are grown in LB broth containing Km and Cm for 16 h at 44°C and analyzed for their plasmid content.

35 One colony is found to carry pPAD5::pMBC2, as verified

BSI 0003 EPN

M 23.06.98

58

by the observation, upon agarose gel-electrophoresis, of fragments of 65, 33, 5.6, 4.7, 3.4, 2.8, 2.1, 1.2, 1.2, 1.0 and 0.4 kb after digestion of the plasmid with *HindIII*.

5

#### Example 69

##### Construction of *E. coli* K12 C600/pPAD6

Several colonies of *E. coli* C600/pPAD5::pMBC2, prepared as described in Example 68, are grown individually in LB broth containing Km for 24 h at 30°C, diluted 1:100 and incubated for further 8 h at 44°C. Appropriate dilutions are plated. Few of the resulting Km<sup>R</sup> Cm<sup>S</sup> Tc<sup>S</sup> colonies are analyzed for their plasmid content. One colony is found to carry pPAD6, as verified by the observation, upon agarose gel-electrophoresis, of fragments of 47, 46, 8.1, 4.6, 2.2, 0.5 and 0.1 kb after digestion of the plasmid with *EcoRI*. After digestion with *DraI* and resolution by PFGE, pPAD6 yields fragments of 102.0, 4.2 and 0.6 kb.

20

Although the present invention is described in the Examples listed above in terms of preferred embodiments, they are not to be regarded as limiting the scope of the invention. The above Examples serve to illustrate the principles and methodologies for assembling pre-existing overlapping segments of DNA into ESAC vectors.

It will occur to those skilled in the art that the GE2270 cluster can be assembled using A-B-C-D fragments other than those specified in the Examples. Any A fragment, such that no biosynthesis genes are present to its left (using the orientation of Fig. 9) can be used for assembling the cluster. Similarly, any D fragment, such that no biosynthesis genes are present to its right (using the orientation of Fig. 9) can also be used. Furthermore, any fragments B common to pRP16

BSI 0003 EPN

M 23.06.98

59

and pRP31, and any fragment C common to pRP31 and pRP58 can be also utilized. It will also occur to those skilled in the art that other methods for obtaining these fragments, such as use of different segments from the cluster of Fig. 9, of different restriction endonucleases, or of the PCR, can be used for achieving equivalent results. It will also occur to those skilled in the art that intermediate vectors, other than the pUC- series used in the above Examples, can be used for subcloning fragments A through D, and that the use of these intermediate vectors is merely instrumental to the transfer of the fragment pairs into the *ts* vector. Some of the fragment pairs could therefore be cloned directly into a *ts* vector.

It will also occur to those skilled in the art that cosmids other than pRP16, pRP31 and pRP58 can be used to achieve equivalent results, provided that they encompass the entire GE2270 gene cluster and they have overlapping segments. It will also occur to those skilled in the art that pMAK705, Lorist6 and pPAC-S1, are merely examples of *ts*, cosmid and ESAC vectors, respectively, that can be used to achieve equivalent results. Any of the several cosmid vectors described in the literature, other *ts* replicons derived from pMAK705 or other source, and any of the ESAC vectors other than pPAC-S1, which include the possible vectors described in Section 7.1, can be used for obtaining equivalent results.

Those skilled in the art understand that the purpose of a *ts* replicon is to select for interplasmid cointegrates at the non-permissive temperature. However, cointegrate formation can occur between any two replicons, and cointegrate can be isolated after transformation of suitable hosts with a plasmid preparation made from an *E. coli* cell harboring both replicons. Selection for the antibiotic resistance

BSI 0003 EPN

60

markers from each replicon can lead to the isolation of cointegrates.

Furthermore, it will occur to those skilled in the art that the inclusion of the *tetR* selectable marker between the A-B, B-C and C-D fragment pairs serve solely the scope of recognizing insert exchange after resolution of the interplasmid cointegrate. Selectable markers other than *tetR* can be equally effective, as long as they are not present in the vectors. In addition, it will occur to those skilled in the art that the presence of a selectable marker is not absolutely necessary, as insert exchange can be observed by other methods, such as selective hybridization or PCR. It will also occur to those skilled in the art that different *E. coli* hosts other than those used in the above Examples can be also employed.

It will also occur to those skilled in the art that, as described in Examples 58 through 69, interplasmid insert exchange can be obtained in a sequel independent of the order of the overlapping cosmid clones in the genomic contig. Indeed, the schematic of Fig. 8 illustrates the sequel of interplasmid exchanges A-B, followed by B-C and then by C-D, while Examples 58 through 69 describe the sequel A-B, C-D and last B-C. Furthermore, it will occur to those skilled in the art that technical variations on the methodologies employed here can produced equivalent results. All these variations fall within the scope of the present invention.

It will occur to those skilled in the art that the principles and methodologies described in Sections 7.2.1 and 7.2.2 are not mutually exclusive. For example, a construct equivalent to pPAD6 can be directly isolated by subjecting the producer strain *P. rosea* ATCC 53733 to the principles and methodologies

BSI 0003 EPN

61

described in Section 7.2.1. Similarly, selected  
cosmids from the described *S. coelicolor* library  
(Redenbach et al., 1996, Mol. Microbiol. 21:77-96) can  
be used for assembling a large chromosomal segment into  
5 pPAC-S1, following the principles and methodologies  
described in Section 7.2.2. Furthermore, it will occur  
to those skilled in the art that the principles and  
methodologies of Section 7.2.1 and 7.2.2 can complement  
each other. For example, after constructing an ESAC  
10 library of *P. rosea* DNA, insert from an individual  
ESAC clone can be enlarged by applying the principles  
and methodologies of Section 7.2.2, using, for example,  
cosmids overlapping that ESAC clone.

Those skilled in the art understand also that the  
15 principles and methodologies described in Sections  
7.2.1 and 7.2.2 and illustrated in schematic form in  
Fig. 3 and Fig. 4, respectively, are general enough  
that they can be applied to other strains and clusters  
responsible for the biosynthesis of different natural  
20 products. Methods for preparing high molecular weight  
DNA, for constructing and propagating in *E. coli* an  
ESAC library can be developed from the principles and  
methodologies described in Examples 12 through 17. The  
principles and methodologies described in the Examples  
25 of Section 7.2.2 can be easily extended to other  
actinomycetes. Methods for preparing the appropriate  
combinations of fragment pairs to yield the starting  
plasmids described in Fig. 6, can be developed  
following the principles and methodologies described in  
30 Examples 19 through 51; methods for assembling the  
entire cluster into an ESAC vector can be developed  
following the principles and methodologies described in  
Examples 52 through 69. In order to illustrate how the  
principles and methodologies described in Section 7.2  
35 can be extended to another actinomycete strain  
producing a different natural product, the present

BSI 0003 EPN

11.23.05.99

62

invention describes hereafter the isolation and construction of an ESAC clone carrying the entire rapamycin biosynthesis gene cluster from the producer strain *Streptomyces hygroscopicus* ATCC 29257. The cluster has been described and is fully contained within three overlapping cosmids designated cos58, cos25 and cos2 (Schwecke et al., 1995, Proc. Natl. Acad. Sci. USA 92:7839-7843). The Examples reported herein describe the construction of the starting plasmids, equivalent to those reported in Fig. 6, that can be used to reassemble the rapamycin biosynthesis gene cluster according to the scheme of Fig. 8.

#### Example 70

##### Preparation of fragment A

Primers 5'-GAATTCGGTACCAGCCGACGGCGA-3' and 5'-GGATCCCTGTTCCACC-AGCGCACC-3' are used to amplify a 1.2 kb fragment from cos58. The fragment is digested with *EcoRI* + *BamHI*.

#### Example 71

##### Preparation of fragment B

Primers 5'-GGATCCAGGAAGCCCTGTGCTGTC-3' and 5'-TCTAGACCGTCGTCGG-TGGTTCTC-3' are used to amplify a 1.2 kb fragment from cos58. The fragment is digested with *BamHI* + *XbaI*.

#### Example 72

##### Construction of plasmid pUR1

Fragment A, prepared as described in Example 71, and fragment C, prepared as described in Example 76, are ligated to pUC18 digested with *EcoRI* + *XbaI*. The resulting mixture contains the desired plasmid pUR1.

BSI 0003 EPN

M 23.06.98

63

Example 73Construction of *E. coli* K12 DH1/ pUR1

Approximately 10 ng of plasmid pUR3, prepared as described in Example 72, are used to transform *E. coli* DH1 and a few of the resulting Ap<sup>R</sup> colonies that appear on the LB-agar plates are analyzed for their plasmid content. One colony is found to carry pUR1, as verified by the observation, upon agarose gel-electrophoresis, of fragments of 2.7 and 2.4 kb after digestion of the plasmid with *EcoRI* + *XbaI*.

Example 74Construction of plasmid pUR2

Plasmid pUR1, prepared as described in Example 73 and previously digested with *BamHI*, and the *tetR* fragment, prepared as described in Example 31, are treated with T4 DNA Polymerase and DNA ligase. The ligation mixture contains the desired plasmid pUR2.

Example 75Construction of *E. coli* K12 DH1/ pUR2

Approximately 10 ng of plasmid pUR4, prepared as described in Example 74, are used to transform *E. coli* DH1 and a few of the resulting Tc<sup>R</sup> Ap<sup>R</sup> colonies that appear on the LB-agar plates are analyzed for their plasmid content. One colony is found to carry pUR2, as verified by the observation, upon agarose gel-electrophoresis, of fragments of 3.9 and 2.8 kb after digestion of the plasmid with *HindIII*.

Example 76Preparation of fragment C

Primers 5'-TGTAGAGGTCAAGATCCGGGGCAT-3' and 5'-CTGCAGGACAGCGCC-CTTGAGGTG-3' are used to amplify a 1.2 kb fragment from cos25. The amplified fragment is digested with *XbaI* and *PstI*.

BSI 0003 EPN

64

Example 77Construction of the plasmid pUR3

Fragment B, prepared as described in Example 70, and  
5 fragment C, prepared as described in Example 75, are  
ligated with pUC18 digested with *Bam*HI + *Pst*I. The  
ligation mixture contains the desired plasmid pUR3.

Example 78Construction of *E. coli* K12 DH1/ pUR3

10 Approximately 10 ng of plasmid pUR2, prepared as  
described in Example 77, are used to transform *E. coli*  
DH1 and a few of the resulting Ap<sup>R</sup> colonies that appear  
on the LB-agar plates are analyzed for their plasmid  
15 content. One colony is found to carry pUR3, as verified  
by the observation, upon agarose gel-electrophoresis,  
of fragments of 2.7 and 2.4 kb after digestion of the  
plasmid with *Bam*HI + *Pst*I.

20

Example 79Construction of the plasmid pUR4

Plasmid pUR3, prepared as described in Example 78 and  
previously digested with *Xba*I, and the *tet*R fragment,  
prepared as described in Example 31, are treated with  
25 T4 DNA Polymerase and DNA ligase. The ligation mixture  
contains the desired plasmid pUR4.

Example 80Construction of *E. coli* K12 DH1/ pUR4

30 Approximately 10 ng of plasmid pUR3, prepared as  
described in Example 79, are used to transform *E. coli*  
DH1 and a few of the resulting Tc<sup>R</sup> Ap<sup>R</sup> colonies that  
appear on the LB-agar plates are analyzed for their  
plasmid content. One colony is found to carry pUR4, as  
35 verified by the observation, upon agarose gel-  
electrophoresis, of fragments of 3.9 and 2.8 kb after

BSI 0003 EPN

M 20.06.98

65

digestion of the plasmid with *Hind*III.

#### Example 81

##### Preparation of fragment D

- 5 Primers 5'-CTGCAGGCGACGAAGAGGGGC-3' and 5'-  
AAGCTTAGCGCGACCGGG-GCGGT-3' are used to amplify a 0.9  
kb fragment from cos2. The amplified fragment is  
digested with *Pst*I and *Hind*III.

10

#### Example 82

##### Construction of the plasmid pUR5

- Fragment C, prepared as described in Example 76, and  
fragment D, prepared as described in Example 81, are  
ligated with pUC18 cut with *Xba*I + *Hind*III. The  
15 ligation mixture contains the desired plasmid pUR5.

#### Example 83

##### Construction of *E. coli* K12 DH1/ pUR5

- Approximately 10 ng of plasmid pUR5, prepared as  
20 described in Example 82, are used to transform *E. coli*  
DH1 and a few of the resulting Ap<sup>R</sup> colonies that appear  
on the LB-agar plates are analyzed for their plasmid  
content. One colony is found to carry pUR5, as verified  
by the observation, upon agarose gel-electrophoresis,  
25 of fragments of 2.7 and 2.1 kb after digestion of the  
plasmid with *Xba*I + *Hind*III.

#### Example 84

##### Construction of the plasmid pUR6

- 30 Plasmid pUR5, prepared as described in Example 83 and  
digested with *Pst*I, and the *tetR* fragment, prepared as  
described in Example 31, are treated with T4 DNA  
Polymerase and DNA ligase. The ligation mixture  
contains the desired plasmid pUR6.

35

BSI 0003 EPN

23.06.98

66

Example 85Construction of *E. coli* K12 DH1/pUR6

Approximately 10 ng of plasmid pUR6, prepared as described in Example 84, are used to transform *E. coli* DH1 and a few of the resulting Tc<sup>R</sup> Ap<sup>R</sup> colonies that appear on the LB-agar plates are analyzed for their plasmid content. One colony is found to carry pUR6, as verified by the observation, upon agarose gel-electrophoresis, of fragments of 3.9 and 2.5 kb after digestion of the plasmid with *Hind*III.

Example 86Construction of plasmid pUR7

The 4.0 kb *Eco*RI-*Xba*I fragment obtained from plasmid pUR2, prepared as described in Example 75, is ligated to the 4.8 kb *Eco*RI-*Xba*I fragment obtained from plasmid pUR5, prepared as described in Example 83. The ligation mixture contains the desired plasmid pUR7.

Example 87Construction of *E. coli* K12 DH1/pUR7

Approximately 10 ng of plasmid pUR7, prepared as described in Example 86, are used to transform *E. coli* DH1 and a few of the resulting Tc<sup>R</sup> Ap<sup>R</sup> colonies that appear on the LB-agar plates are analyzed for their plasmid content. One colony is found to carry pUR7, as verified by the observation, upon agarose gel-electrophoresis, of fragments of 4.9 and 3.9 kb after digestion of the plasmid with *Hind*III.

Example 88Construction of plasmid pMR1

The 4.0 Kb *Eco*RI-*Xba*I fragment obtained from plasmid pUR2, prepared as described in Example 75, is treated with DNA Polymerase to fill-in the ends and ligated to pMAK705 digested with *Hinc*II. The ligation mixture

BSI 0003 EPN

11 25 06 98

67

contains the desired plasmid pMR1.

#### Example 89

##### Construction of *E. coli* K12 DH1/pMR1

- 5 Approximately 10 ng of plasmid pMR1, prepared as described in Example 88, are used to transform *E. coli* DH1 and a few of the resulting  $\text{Cm}^R$   $\text{Tc}^R$  colonies that appear on the LB-agar plates are analyzed for their plasmid content. One colony is found to carry pMR1, as
- 10 verified by the observation, upon agarose gel-electrophoresis, of fragments of 6.7 and 2.8 kb after digestion of the plasmid with *HindIII*.

#### Example 90

##### Construction of plasmid pMR2

- 15 The 4.0 kb *EcoRI*-*PstI* fragment obtained from plasmid pUR4, prepared as described in Example 80, after filling-in of the resulting ends, is ligated to pMAK705 digested with *HincII*. The ligation mixture contains the
- 20 desired plasmid pMR2.

#### Example 91

##### Construction of *E. coli* K12 DH1/pMR2

- 25 Approximately 10 ng of plasmid pMR2, prepared as described in Example 90, are used to transform *E. coli* DH1 and a few of the resulting  $\text{Cm}^R$   $\text{Tc}^R$  colonies that appear on the LB-agar plates are analyzed for their plasmid content. One colony is found to carry pMR2, as
- 30 verified by the observation, upon agarose gel-electrophoresis, of fragments of 6.7 and 2.8 kb after digestion of the plasmid with *HindIII*.

#### Example 92

##### Construction of plasmid pMR3

- 35 The 3.7 kb *EcoRI*-*NdeI* fragment from plasmid pUR6, prepared as described in Example 85, obtained after

BSI 0003 EPN

M 23.05.98

68

partial digestion with *EcoRI* and complete digestion with *NdeI*, after filling-in of the resulting ends, is ligated to pMAK705 previously digested with *HincII*. The ligation mixture contains the desired plasmid pMR3.

5

#### Example 93

##### Construction of *E.coli* K12 DH1/pMR3

Approximately 10 ng of plasmid pMR3, prepared as described in Example 92, are used to transform *E. coli* DH1 and a few of the resulting  $\text{Cm}^R$   $\text{Tc}^R$  colonies that appear on the LB-agar plates are analyzed for their plasmid content. One colony is found to carry pMR3, as verified by the observation, upon agarose gel-electrophoresis, of fragments of 6.7 and 2.5 kb after digestion of the plasmid with *HindIII*.

15

#### Example 94

##### Construction of plasmid pPR1

The 6.1 kb *EcoRI*-*NdeI* fragment obtained from plasmid pUR7, prepared as described in Example 87, is ligated to pPAC-S1, prepared as described in Example 11, previously digested with *ScaI*. The ligation mixture contains the desired plasmid pPR1.

20

#### Example 95

##### Construction of *E.coli* K12 DH1/pPR1

Approximately 10 ng of plasmid pPR1, prepared as described in Example 94, are used to transform *E. coli* DH1 and a few of the resulting  $\text{Km}^R$   $\text{Tc}^R$  colonies that appear on the LB-agar plates are analyzed for their plasmid content. One colony is found to carry pPR1, as verified by the observation, upon agarose gel-electrophoresis, of fragments of 19.7, 4.7 and 1.4 kb after digestion of the plasmid with *BamHI*.

30

BSI 0003 EPN

M 25.05.98

69

The Examples reported above describe the principle and methodologies for constructing plasmids equivalent to those illustrated in Fig. 6, that can be used to assemble the rapamycin gene cluster into the ESAC

5 vector pPAC-S1. It will occur to those skilled in the art that the principles and methodologies illustrated in Fig. 7 and described in Examples 52 through 69 can be applied to the rapamycin gene cluster, employing the pMAK705 derivatives constructed in Examples 89, 91 and

10 93, and cosmids cos58, cos25 and cos2 (Schwecke et al., 1995, Proc. Natl. Acad. Sci. USA 92:7839-7843). Thus, as described in Examples 52 through 57 for the assemblage of the GE2270 gene cluster, pMR1, prepared as described in Example 89, and cos58 (Schwecke et al.,

15 1995, Proc. Natl. Acad. Sci. USA 92:7839-7843) substitute pMAB1 and pRP16, respectively, in Example 52; pMR2, prepared as described in Example 91, and cos25 (Schwecke et al., 1995, Proc. Natl. Acad. Sci. USA 92:7839-7843) substitute pMBC1 and pRP31,

20 respectively, in Example 54; and pMR3, prepared as described in Example 93, and cos2 (Schwecke et al., 1995, Proc. Natl. Acad. Sci. USA 92:7839-7843) substitute pMCD1 and pRP58, respectively, in Example 56. Those skilled in the art understand that the

25 principles and methodologies illustrated in Fig. 8 can be applied employing pPR1, prepared as described in Example 95, and the pMAK705 derivatives obtained after insert exchange between cos58, cos 25 and cos 2, and pMR1, pMR2 and pMR3, respectively. In analogy with

30 Examples 58 through 69, the equivalent constructs pPR2 through pPR6 can be generated.

It will occur to those skilled in the art that, although illustrated in Fig. 5 through 8 by three overlapping clones and described in the Examples 58

35 through 69 by the use of five rounds of interplasmid insert exchange, the principles and methodologies

BSI 0003 EPN

23.06.98

70

described in this section of the present invention can be extended to a different number of overlapping clones. If  $n$  is the number of overlapping clones that encompass the desired genomic segment,  $n$  will also be the number of homologous recombination rounds that introduce cluster DNA into the ESAC vector. If an  $Ab^R$  marker is used to facilitate monitoring insert exchange, the total number of rounds of homologous recombination will be equal to  $2n - 1$ , when  $n$  is the number of overlapping clones. Interplasmid homologous recombination has been described to introduce large DNA segments into a desired vector (O'Connor et al., 1989, Science 244:1307-1312; Kao et al., 1994, Science 265:509-512) or to target a smaller segment into a large episome (Yang et al., 1997, Nature Biotechnol. 15:859-865). However, it was not anticipated that these procedures could be applied iteratively for the precise reconstruction of very large DNA segments.

### 7.3 Identification of positive clones

The principles and methodologies described in Section 7.2 for obtaining an entire gene cluster in an ESAC vector rely on the identification of the desired genomic segment. When using the principles and methodologies described in Section 7.2.1, the desired clones are identified by screening an ESAC library with one of the possible strategies described below. When using the principles and methodologies described in Section 7.2.2, the desired clones are identified in a genomic library, such as a cosmid library, with one of the possible strategies described below, and then assembled into ESAC. The principles and methodologies for identifying the genes responsible for the biosynthesis of natural products are well described in the literature and are reported here solely to illustrate the fact that they represent a necessary

BSI 0003 EPN

M 25.06.98

71

step in the overall scope of the present invention.

The genes involved in the biosynthesis of natural products in actinomycetes are invariably found as gene clusters in the chromosome of the producing organism, often associated with one or more resistance determinants. Consequently, identifying one gene allows ready access to all the others. One or more genes responsible for the biosynthesis of a natural product could have been described, or the entire cluster could be known. Several biosynthesis clusters from actinomycetes have been reported (Peshcke et al., 1995, *Mol. Microbiol.* 16:1137-1156; Vining and Stuttard, 1995, *Genetics and Biochemistry of Antibiotic Formation*, Butterworth-Heinemann, Boston, CT, USA; Schwecke et al., 1995, *Proc. Natl. Acad. Sci. USA* 92:7839-7843; August et al., 1998, *Chem. Biol.* 5:69-79; van Wageningen et al., 1998, *Chem. Biol.* 5:155-162). Other clusters are likely to be described in the future. Suitable probes from the cluster extremities can be derived from published clusters, when available.

If no biosynthesis genes are known, different strategies for identifying them can be applied. These strategies are well described in the literature and have been widely employed. They are summarized below.

One possible strategy involves the isolation of the resistance gene(s) after cloning it in a heterologous host that is sensitive to that natural product (which is, in this case, an antibiotic: for example, Stanzak et al., 1986, *Bio/Technol.* 4:229-232). Another possible strategy for identifying biosynthesis genes is based on reverse genetics: a particular biosynthetic enzyme is purified, and from its partial protein sequence(s) the corresponding gene is isolated via PCR or hybridization: for example, Fishman et al., 1987, *Proc. Natl. Acad. Sci. USA* 84:8248-8252. Another approach relies on the complementation of mutants blocked in one

BSI 0003 EPN

11.23.05.98

72

or more steps of the biosynthesis, after introduction of a DNA library constructed in a suitable vector into the wild type strain: for example, Malpartida and Hopwood, 1984, *Nature* 309:462-464. Another approach involves the construction of an expression library in a suitable vector in an appropriate host, where the gene product is sought after using specific antibodies or looking for a particular enzymatic activity (for example, Jones and Hopwood, 1984, *J. Biol. Chem.* 259:14151-14157).

Other possible approaches make use of heterologous probes derived from biosynthesis, resistance and regulatory genes. Natural products can be broadly grouped into classes according to their biosynthetic origin. For many of these major classes suitable probes are available. For example, genes encoding aromatic or modular polyketide synthases can be effectively identified through the use of heterologous hybridization probes (Malpartida et al., 1987, *Nature* 325:818-821; Schwecke et al., 1995, *Proc. Natl. Acad. Sci. USA* 92:7839-7843); suitable probes have been reported for peptide synthetase genes (Turgay and Marahiel, 1994, *Pept. Res.* 7:238-241); for genes involved in the formation or attachment of modified sugars (Decker et al., 1996, *FEMS Microbiol. Lett.* 141:195-201). As our understanding of the genetics of natural product biosynthesis increases, other heterologous probes will become available.

The size of clusters can be estimated from those of known clusters involved in the synthesis of structurally similar natural products. For examples, synthesis of macrolides is expected to require gene clusters in the 60-70 kb range (Katz and Donadio, 1993, *Annu. Rev. Microbiol.* 47:875-912; Kuhstoss et al., 1996, *Gene* 183:231-236); synthesis of glycopeptides, clusters in the 70 kb range (van Wageningen et al.,

BSI 0003 EPN

73

Sci. USA 92:7839-7843); suitable probes have been reported for peptide synthetase genes (Turgay and Marahiel, 1994, Pept. Res. 7:238-241); for genes involved in the formation or attachment of modified  
5 sugars (Decker et al., 1996, FEMS Microbiol. Lett. 141:195-201). As our understanding of the genetics of natural product biosynthesis increases, other heterologous probes will become available.

The size of clusters can be estimated from those  
10 of known clusters involved in the synthesis of structurally similar natural products. For examples, synthesis of macrolides is expected to require gene clusters in the 60-70 kb range (Katz and Donadio, 1993, Annu. Rev. Microbiol. 47:875-912; Kuhstoss et al.,  
15 1996, Gene 183:231-236); synthesis of glycopeptides, clusters in the 70 kb range (van Wageningen et al., 1998, Chem. Biol. 5:155-162). In instances where no clusters have been described for the same structural class of natural products, the size of the relevant  
20 cluster can be estimated from considerations about its biosynthesis route, which can often be derived from analogy to other natural products. Once the desired cluster has been identified, its extent can be established by analysis of the DNA sequence of the  
25 cloned cluster or of parts thereof. Comparison of the DNA sequence to databases can allow the identification of the likely borders of the gene cluster.

Employing any of the above mentioned approaches, the clones containing the desired gene cluster can be  
30 identified in an ESAC library (prepared according to the principles and methodologies of Section 7.2.1), or assembled into an ESAC vector (according to the principles and methodologies of Section 7.22). Since the amount of genetic information that can be  
35 introduced into ESAC vectors has no defined upper limit, DNA inserts larger than what may be strictly

BSI 0003 EPN

11 23 06 98

74

necessary to direct the biosynthesis of the natural product may be inserted into an ESAC vector.

#### 7.4 Transformation of a *Streptomyces* host

- 5 Once the desired gene cluster has been introduced in an ESAC vector, according to the principles and methodologies of Section 7.2, one or a few ESAC clones are introduced into a suitable *Streptomyces* host. This is accomplished by employing published procedures for
- 10 transformation of *Streptomyces*. Only minor modifications from established procedures (Hopwood et al., 1985, *Genetic Manipulation of Streptomyces: A Laboratory Manual*, The John Innes Foundation, Norwich, UK) are required for obtaining a sufficient number of
- 15 transformants. Because transformations are performed with single, purified ESAC clones, transformation efficiencies do not need to be particularly high. The Examples reported below illustrate the principles and methodologies for introducing ESAC clones into *S.*
- 20 *lividans*. They serve to describe the present invention and are not meant to restrict its scope. *Streptomyces* transformants are selected for  $\text{Th}^R$ , specified by the *tsr* marker present in the ESAC vector. Since the incoming DNA is incapable of self-replication in
- 25 *Streptomyces*, site-specific integration occurs at the chromosomal *attB* site, mediated by the *int-attP* function specified by the ESAC vector. That integration has occurred at the proper site can be verified by Southern hybridization or by PFGE analysis
- 30 of the transformants. Fig. 11 illustrates a PFGE separation of a *S. lividans* derivative carrying an ESAC clone with a 70 kb insert integrated into its chromosome.

BSI 0003 EPN

75

Example 96Introduction of ESAC clones into *S. lividans* ZX7

A few hundred ng of three individual ESAC clones, prepared as described in Example 17 and carrying  
5 inserts of *S. coelicolor* DNA of 70, 120 and 140 kb (designated ESAC-70, ESAC-120, and ESAC-140, respectively), are used to transform protoplasts of *S. lividans* ZX7. The colonies that appear on the R2YE plates, after overlaying with Th, are analyzed for  
10 their Th<sup>R</sup> by streaking them on fresh R2YE plates.

Example 97Cultivation and preservation of *S. lividans* ZX7/ESAC

Individual colonies of *S. lividans* ZX7 transformants  
15 with the individual ESAC clones, prepared as described in Example 96, are grown for several passages in solid medium without and with Th. Spore suspension, or mycelium prepared after cultivation in JM or YEME medium with Th, are stored at -80°C after addition of  
20 glycerol to 20% (v/v).

Example 98Characterization of *S. lividans* ZX7 attB::ESAC-70

Individual colonies of *S. lividans* ZX7 attB::ESAC-70,  
25 prepared as described in Example 96, are grown in YEME and total genomic DNA is prepared. The DNA is digested with BamHI, resolved by agarose gel-electrophoresis, and transferred onto a membrane. Hybridization to labeled pPAC-S1 DNA, prepared as described in Example  
30 11, reveals the appearance of three bands of approximately 16, 8 and 2.7 kb. PFGE analysis of genomic DNA reveals the disappearance of a 2.5 Mb DraI fragment present in ZX7 and the appearance of two fragments of 1.4 and 1.1 Mb (Fig. 11).

35

BSI 0003 EPN

M 23.05.98

76

Although the present invention is described in the Examples listed above in terms of preferred embodiments, they are not to be regarded as limiting the scope of the invention. The above Examples serve to

5 illustrate the principles and methodologies for introducing ESAC clones, prepared as described in Section 7.2, into *S. lividans*, for the cultivation of the resulting transformants, and for the characterization of their genotype. The above Examples

10 serve to illustrate the principles and methodologies for the transformation of *S. lividans* with ESAC clones carrying DNA inserts from a different species. It will occur to those skilled in the art that additional ESAC clones, either containing different inserts of *S.*

15 *coelicolor* DNA, prepared as described in Example 17, or carrying DNA inserts from other actinomycetes, prepared according to the principles and methodologies of Section 7.2, can be used to transform *S. lividans* ZX7.

It will also occur to those skilled in the art that

20 other *S. lividans* strains can be equally used as hosts for transformation with ESAC clones. Furthermore, phage  $\Phi$ C31 can lysogenize other *Streptomyces* species, in addition to *S. lividans*. These include but are not limited to the species reported in Table 2.

25 Furthermore, it will occur to those skilled in the art that a  $\Phi$ C31 *attB* site could be engineered into *Streptomyces* species or other actinomycetes that are not naturally lysogenized by phage  $\Phi$ C31. Therefore, any ESAC clone, prepared according to the principles

30 and methodologies of Section 7.2, and any natural or engineered actinomycete host, fall within the scope of the present invention.

It will occur to those skilled in the art that alternative methods for introducing DNA into *S.*

35 *lividans* can be employed. These include but are not limited to electroporation (MacNeil, 1989, FEMS

BSI 0003 EPN

M 23.06.98

77

Microbiol. Lett. 42:239-244) and conjugation from *E. coli* (Mazodier et al., 1989, J. Bacteriol. 171:3583-3585). It will also occur to those skilled in the art that alternative media and growth conditions can be employed for cultivating *S. lividans* transformants, and that they can be analyzed by different methods than those described above. Technical variations on the methodologies described above can produced equivalent results. All these variations fall within the scope of the present invention.

BSI 0003 EPN

N 23.08.98

78

---

Table 2

List of exemplary species of *Streptomyces* and other  
5 genera of *Actinomycetales* allowing attP-mediated  
integration of  $\Phi$ C31 (Hopwood et al., 1985, Genetic  
Manipulation of *Streptomyces*: A Laboratory Manual, The  
John Innes Foundation, Norwich, UK; Lomovskaya et al.,  
1997, Microbiol. 143:875-883; Kuhstoss et al., 1991,  
10 Gene 97:143-146; Soldatova et al., 1994, Antibiot.  
Khimioter. 39:3-7).

---

*Streptomyces coelicolor*  
15 *Streptomyces lividans*  
*Streptomyces hygroscopicus*  
*Streptomyces bambergiensis*  
*Streptomyces ambofaciens*  
*Streptomyces griseofuscus*  
20 *Streptomyces lipmanii*  
*Streptomyces thermotolerans*  
*Streptomyces clavuligerus*  
*Streptomyces fradiae*  
25 *Saccharopolyspora spinosa*

---

BSI 0003 EPN

M 23 06 98

79

### 7.5 Growth of the recombinant *Streptomyces* and antibiotic production

When an ESAC clone, introduced into a production host according to the principles and methodologies described in Section 7.4, carries the entire biosynthesis gene cluster derived from a donor organism, the recombinant strain will produce the relevant natural product. Naive actinomycete hosts have been shown to produce the appropriate natural product or its intermediate(s) when the relevant DNA was introduced into them (Malpartida and Hopwood, 1984, *Nature* 309:462-464; Hong et al., 1997, *J. Bacteriol.* 179:470-476; Kao et al., 1994, *Science* 265:509-512). While the examples reported thus far represent special cases (relatively small clusters, careful engineering of selected biosynthesis genes), transformants of *Streptomyces* and other actinomycete species with the relevant biosynthesis clusters are expected to produce the corresponding natural product. The recombinant production hosts are cultivated in a suitable medium and the presence of the relevant metabolites is determined following appropriate procedures. The Examples reported below describe the production of the antibiotic GE2270 by a *S. lividans* ZX7 transformant obtained after introduction of the GE2270 biosynthesis gene cluster cloned into an ESAC vector. The Examples reported below illustrate the principles and methodologies for achieving production of a secondary metabolite by a recombinant *S. lividans*. They serve to describe the present invention and are not meant to restrict its scope.

#### Example 99

##### Construction *S. lividans* ZX7 attB::pPAD6

A few hundred ng of pPAD6, prepared as described in Example 69, are used to transform protoplasts of *S.*

BSI 0003 EPN

M 23.06.98

80

*lividans* ZX7. The colonies that appear on the R2YE plates, after overlaying with Th, are analyzed for their Th<sup>R</sup> by streaking them on fresh R2YE plates.

5

Example 100Isolation of GE2270 from *S. lividans* ZX7 attB::pPAD6

A single colony of *S. lividans* ZX7 attB::pPAD6, prepared as described in Example 99, is inoculated into JM medium containing Th and grown for 48 h at 30°C. The culture is diluted 1:10 with 300 mL of fresh JM medium containing Th and grown for further 72 h. The antibiotic is extracted from the culture broth and from the harvested mycelium. Identification of GE2270 is performed by MS and NMR.

15

Although the present invention is described in the Examples listed above in terms of preferred embodiments, they are not to be regarded as limiting the scope of the invention. The above Examples serve to illustrate the principles and methodologies for achieving production of GE2270 after introduction of pPAD6 into *S. lividans*. It will occur to those skilled in the art that alternative methods for introducing DNA into *S. lividans*, alternative media and growth conditions for the transformants, and alternative methods for producing, isolating and analyzing GE2270 can be employed. Technical variations on the methodologies described above can produced equivalent results and fall within the scope of the present invention.

30

It will occur to those skilled in the art that ESAC clones, containing the relevant biosynthesis cluster derived from different donor organisms, can be used to transform *S. lividans*. The resulting transformants will produce the corresponding natural product. For example, an ESAC clone carrying the

35

BSI 0003 EPN

81

rapamycin cluster, prepared according to the principles of Section 7.2, can be used to transform *S. lividans* and rapamycin can be produced by the resulting recombinant strain. Furthermore, it will occur to those skilled in the art that other *Streptomyces* or actinomycete strains that naturally contain or have been engineered to contain a phage  $\Phi$ C31 attB site, can be used as production hosts for desired natural products. Therefore, any natural product produced after introduction into *S. lividans* of the relevant genes carried on ESAC, falls within the scope of the present invention.

The present invention describes principles and methodologies for optimizing and speeding up the process of lead optimization and development in drug discovery. The present invention can be applied since the early stages of drug discovery as briefly summarized herein. A natural product produced by a donor organism has interesting properties, such as antibacterial, antifungal, antitumor, antihelmintic, immunosuppressive, herbicidal or other pharmacological activity. The potential is seen for increasing the productivity of the producing organism, and/or for improving the biological or physico-chemical properties of said natural product after manipulating its structure. The biosynthetic pathway for the natural product is inferred from its chemical structure. This leads to a hypothesis on the genes involved, including the approximate size of the corresponding cluster. A large insert library is constructed in the ESAC vectors described herein using genomic DNA prepared from the donor organism. Through a judicious choice of hybridization probes and PCR primers, the desired cluster is identified in the library. Alternatively, the cluster is assembled into the ESAC vectors described herein from overlapping cosmid clones

BSI 0003 EPN

N 23.08.98

82

identified by hybridization as above. The selected clone(s) are transferred into *S. lividans*, *S. coelicolor* or other suitable strain, and the resulting transformants are evaluated for production of the

5 natural product.

Once production is obtained, the desired genetic, physiological and technological manipulations can be performed on the production host, employing well-developed methodologies. The bioactive molecule is

10 purified from a known host, amid a background of known metabolites. If necessary, *ad hoc* mutations can be conveniently introduced in the production host to eliminate unwanted, interfering products. Because of the deeper knowledge on the physiological processes and

15 regulatory networks for secondary metabolism in the production host compared to the donor organism, targeted approaches to strain improvement, using classical and molecular techniques, can be applied. Furthermore, well-characterized mutant strains can be

20 available for the producing host, and the desired ESAC clone could be easily introduced into different genetic backgrounds. In addition, the biosynthetic pathway can be easily manipulated, because of the availability of the cloned genes and the efficient genetic tools

25 available for the production host. Finally, additional specialized genes or even entire clusters can be introduced into the production host, further expanding the possible applications of the present invention.

Finally, even in a case where the natural product

30 may not be made by the production host after transfer of the relevant cluster, appropriate tools are available to remedy that situation. For example, lack of production of the expected natural product could be due to several possibilities: absence of required

35 gene(s); DNA, gene product or natural product instability; inadequate levels of gene expression or of

BSI 0003 EPN

83

appropriate precursors. In a well-defined production host, each of these possible causes can be accessed by direct experimentation and remedied.

Therefore, the present invention provides

5 significant advantages over the existing process of drug discovery and development, including production. It exploits the fact that the host where the natural product will be produced is an organism commonly used for process development and genetic manipulation, with

10 substantial information available, including safety for handling it.

BSI 0003 EPN

M 23.06.98

84

SUMMARY OF INFORMATION SUPPLIED WITH THE DATA CARRIER  
PATENTIN VERSION 1.30

(Sequence no. 1)

5

## (1) GENERAL INFORMATION

APPLICANT: BIOSEARCH ITALIA S.p.A.  
Via Roberto Lepetit, 34  
21040 GERENZANO (Varese)

10

## (6) BIOLOGICAL AND SEQUENCE INFORMATION

## (3) Source of organism information

15

Molecule type: DNA (genomic)  
Scientific name: Planisbora rosea  
Strain name: ATCC 53733

(5) Sequence editor: SEQUENCE no. 1

20

GGATCCCGAGCACCGACCAGCCGTGGGCGGGGACGAGACACGGGTCTCCCGGAGC  
CTCCCCCGACGACTCCAGCACGGCCAGGCCCGCGGCCTCGACCGGGAAGCGGTAG  
GGCCTGTCGTCCACGGTTGAGCAGGGTGAGCAGTGCCCGGCCGGGATGGTCCGGG  
TCAGCCGAGGCCAGCGCGGCGGCCCGGTTGCTCAG

BSI 0003 EPN

85

SUMMARY OF INFORMATION SUPPLIED WITH THE DATA CARRIER  
PATENTIN VERSION 1.30

(Sequence no. 2)

5

## (1) GENERAL INFORMATION

APPLICANT: BIOSEARCH ITALIA S.p.A.  
Via Roberto Lepetit, 34  
21040 GERENZANO (Varese)

## (6) BIOLOGICAL AND SEQUENCE INFORMATION

## 15 (3) Source of organism information

Molecule type: DNA (genomic)

Scientific name: *Planisbora rosea*

Strain name: ATCC 53733

(5) Sequence editor: SEQUENCE no. 2

20 (6)

CCGGGAGATCCGCCGACGCCGGCGGGCCGTGCACCACGGTCCTCCTGTTCCGGGCG  
ACGGTGAACGGCAGGCAGGTCCACGGTTCCGATTTCCTGCACTTCGACGACGACG  
GCCTCATCGGCGAGCTCACCGTCATGGTCCGGCCGCTGTCGGGAGCGAACGCGCT  
GGCCGAGGCGATGGGCGCCCAGTTCGAACGGATCC

25



BSI 0003 EPN

86

(5)

## CLAIMS

1. A method for transferring the formation of a  
5 natural product from an actinomycete donor organism  
that is the original producer of said natural product  
to a different actinomycete production host, where this  
transfer is achieved through the use of an *E. coli*-  
*Streptomyces* artificial chromosome.

10 2. An actinomycete strain that is constructed from  
a production host after transfer of a cluster from a  
donor organism as recited in claim 1.

15 3. A gene cluster governing the biosynthesis of a  
natural product from a donor organism as recited in  
claim 1 that is carried on an *E. coli*-*Streptomyces*  
artificial chromosome.

20 4. An *E. coli* host that contains an *E. coli*-  
*Streptomyces* artificial chromosome carrying a gene  
cluster governing the biosynthesis of a natural product  
as recited in claim 3 and that can be used as a source  
of DNA for transformation of a production host as  
recited in claim 2.

25 5. An *E. coli*-*Streptomyces* artificial chromosome  
that carries a gene cluster governing the biosynthesis  
of a natural product as recited in claim 3 and that can  
be used for the purpose recited in claim 1.

30 6. A library constructed in an *E. coli*-  
*Streptomyces* artificial chromosome as recited in claim  
5 using DNA prepared from an actinomycete donor  
organism.

7. An *E. coli*-*Streptomyces* artificial chromosomes  
as recited in claim 5 that is the vector pPAC-S1  
described in Fig. 2.

35 8. An *E. coli*-*Streptomyces* artificial chromosomes  
as recited in claim 5 that is the vector pPAC-S2  
described in Fig. 2.

BSI 0003 EPN

87

9. An *E. coli-Streptomyces* artificial chromosome as recited in claim 5 that carries the gene cluster governing the biosynthesis of the antibiotic GE2270.

5 10. An *E. coli-Streptomyces* artificial chromosome as recited in claim 9 that carries the gene cluster governing the biosynthesis of the antibiotic GE2270 from *Planobispora rosea* ATCC 53733.

10 11. An *E. coli-Streptomyces* artificial chromosome as recited in claim 10 that carries the gene cluster reported in Fig. 9.

12. An *E. coli-Streptomyces* artificial chromosome as recited in claim 11 that is the construct pPAD6, useful for transferring the production of GE2270 to an actinomycete production host.

15 13. The GE2270 gene cluster as reported in Fig. 9, useful for increasing the yield of GE2270 and for producing novel derivatives of GE2270.

20 14. The signature sequences of Fig. 10, useful for identifying the GE2270 gene cluster as recited in claim 13.

15. An actinomycete production host as recited in claim 2 that contains the *E. coli-Streptomyces* artificial chromosome as recited in claim 9.

25 16. An actinomycete production host as recited in claim 2 that contains the *E. coli-Streptomyces* artificial chromosome as recited in claim 10.

17. An actinomycete production host as recited in claim 2 that contains the *E. coli-Streptomyces* artificial chromosome as recited in claim 11.

30 18. An actinomycete production host as recited in claim 2 that contains the *E. coli-Streptomyces* artificial chromosome as recited in claim 12.

19. An actinomycete production host as recited in claim 15 that is *Streptomyces lividans*.

35 20. An actinomycete production host as recited in claim 16 that is *Streptomyces lividans*.

BSI 0003 EPN

88

21. An actinomycete production host as recited in claim 17 that is *Streptomyces lividans*.

22. An actinomycete production host as recited in claim 18 that is *Streptomyces lividans*.

5 23. A *Streptomyces lividans* production host as recited in claim 19 that can be transformed into a high producer of the antibiotic GE2270.

24. A *Streptomyces lividans* production host as recited in claim 19 that can be transformed into a  
10 producer of a controlled complex of the antibiotic GE2270.

25. A *Streptomyces lividans* production host as recited in claim 19 that can be transformed into a producer of a derivative of the antibiotic GE2270.

26. An *E. coli* strain that contains the *E. coli*-*Streptomyces* artificial chromosome as recited in claim 9.

27. An *E. coli* strain that contains the *E. coli*-*Streptomyces* artificial chromosome as recited in claim 10.

28. An *E. coli* strain that contains the *E. coli*-*Streptomyces* artificial chromosome as recited in claim 11.

29. An *E. coli* strain that contains the *E. coli*-*Streptomyces* artificial chromosome as recited in claim 12.

30. The *E. coli*-*Streptomyces* artificial chromosome that is pPAD5, useful for constructing the *E. coli*-*Streptomyces* artificial chromosome as recited in claim 12.

31. The *E. coli*-*Streptomyces* artificial chromosome that is pPAD4, useful for constructing the *E. coli*-*Streptomyces* artificial chromosome as recited in claim 30.

32. The *E. coli*-*Streptomyces* artificial chromosome that is pPAD3, useful for constructing the *E. coli*-*Streptomyces* artificial chromosome as recited in claim 31.

33. The *E. coli*-*Streptomyces* artificial chromosome that is pPAD2, useful for constructing the *E. coli*-*Streptomyces* artificial chromosome as recited in claim 32.

34. The *E. coli*-*Streptomyces* artificial chromosome that is pPAD1, useful for constructing the *E. coli*-*Streptomyces*

23-06-1998

EP98111506.6

CLMS

BSI 0003 EPN

M 23.06.98

89

artificial chromosome as recited in claim 33.

BSI 0003 EPN

90

1/11

6

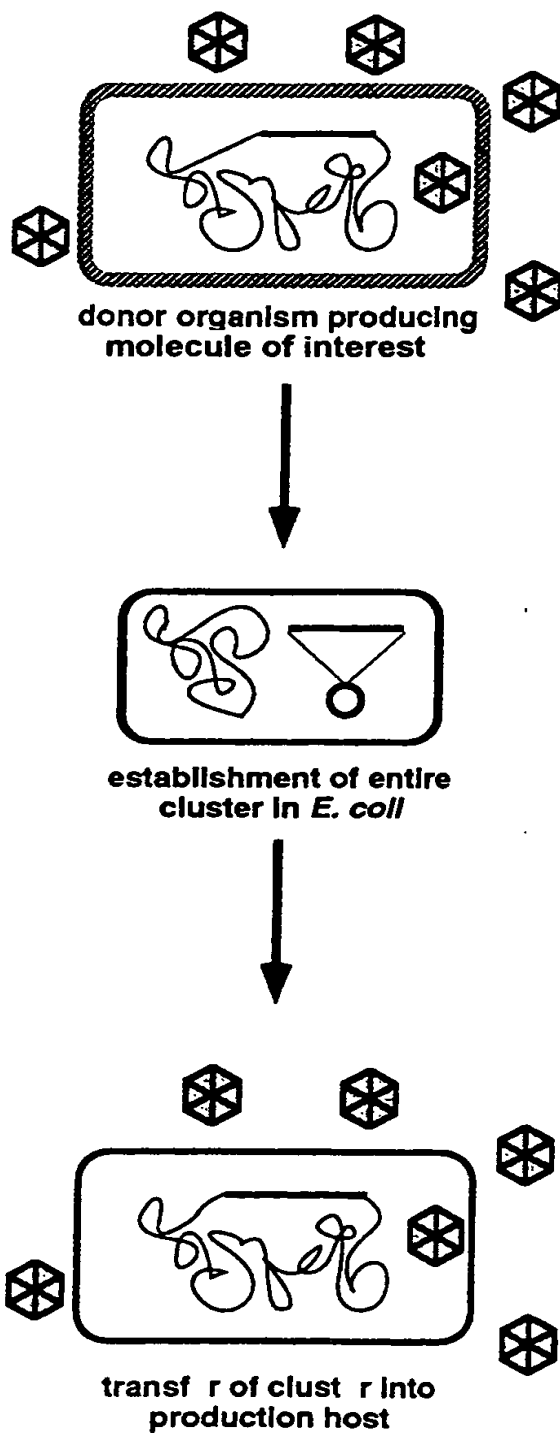


Figure 1

BSI 0003 EPN

23-06-98

91

2/11

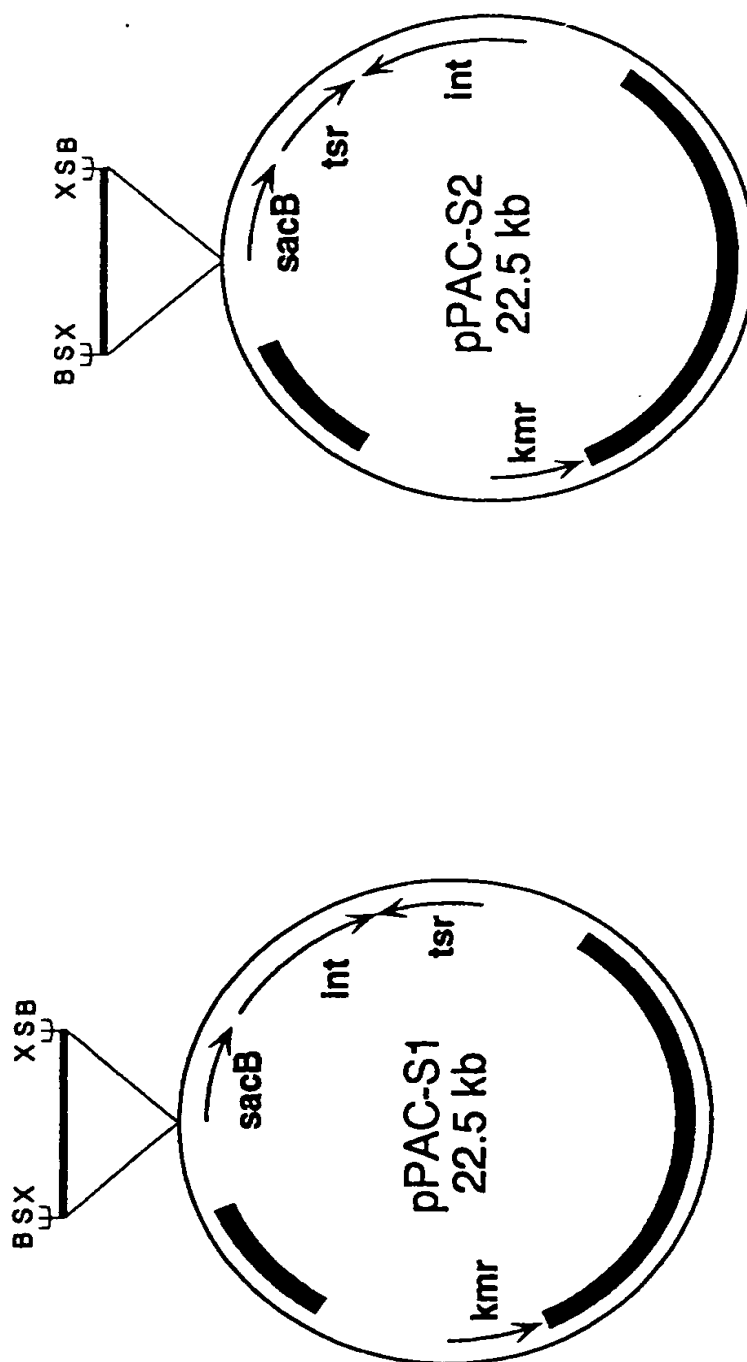
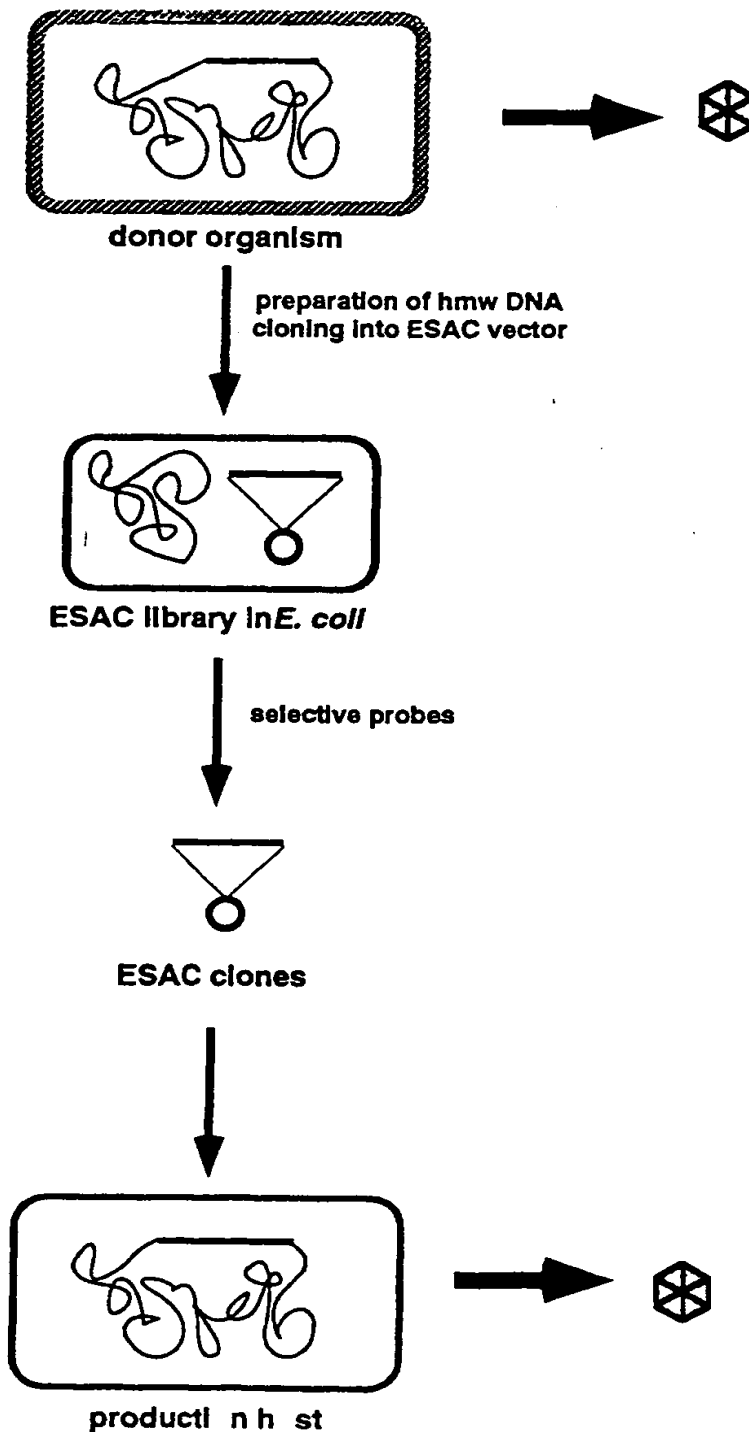


Figure 2

BSI 0003 EPN

23.06.98

92  
3/11

BSI 0003 EPN

M 23.08.98

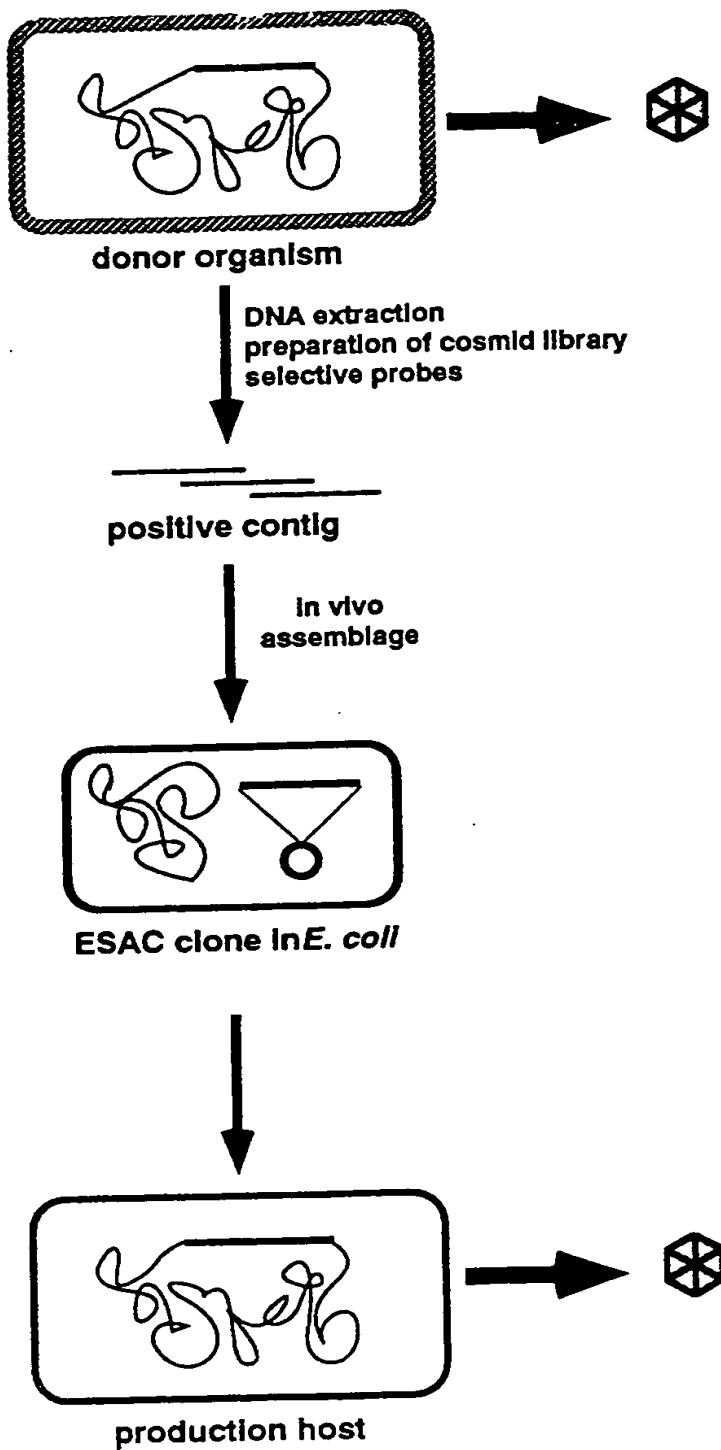
93  
4/11

Figure 4

BSI 0003 EPN

23.06.98

94  
5/11

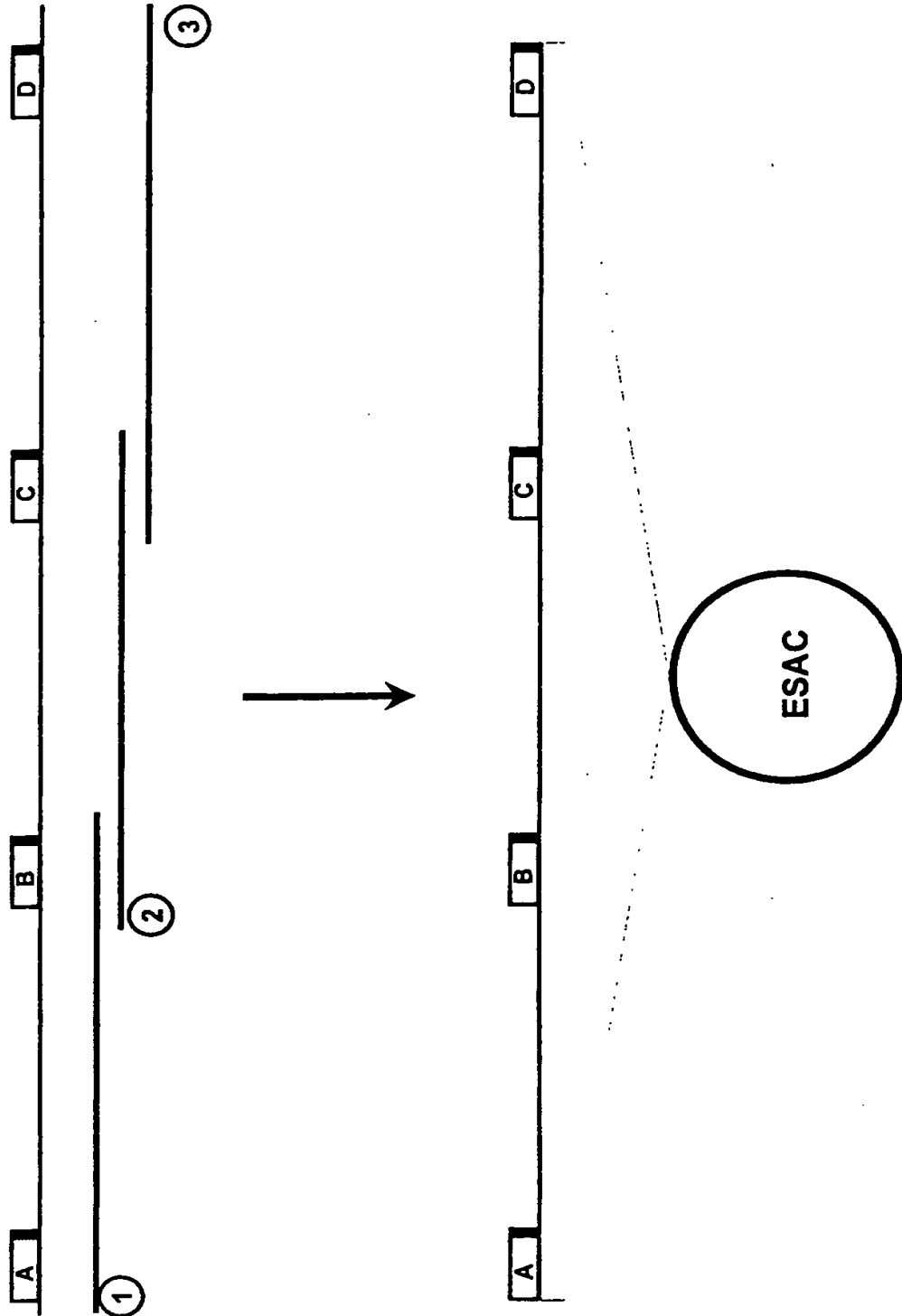


Figure 5

23.06.98

95

6/11

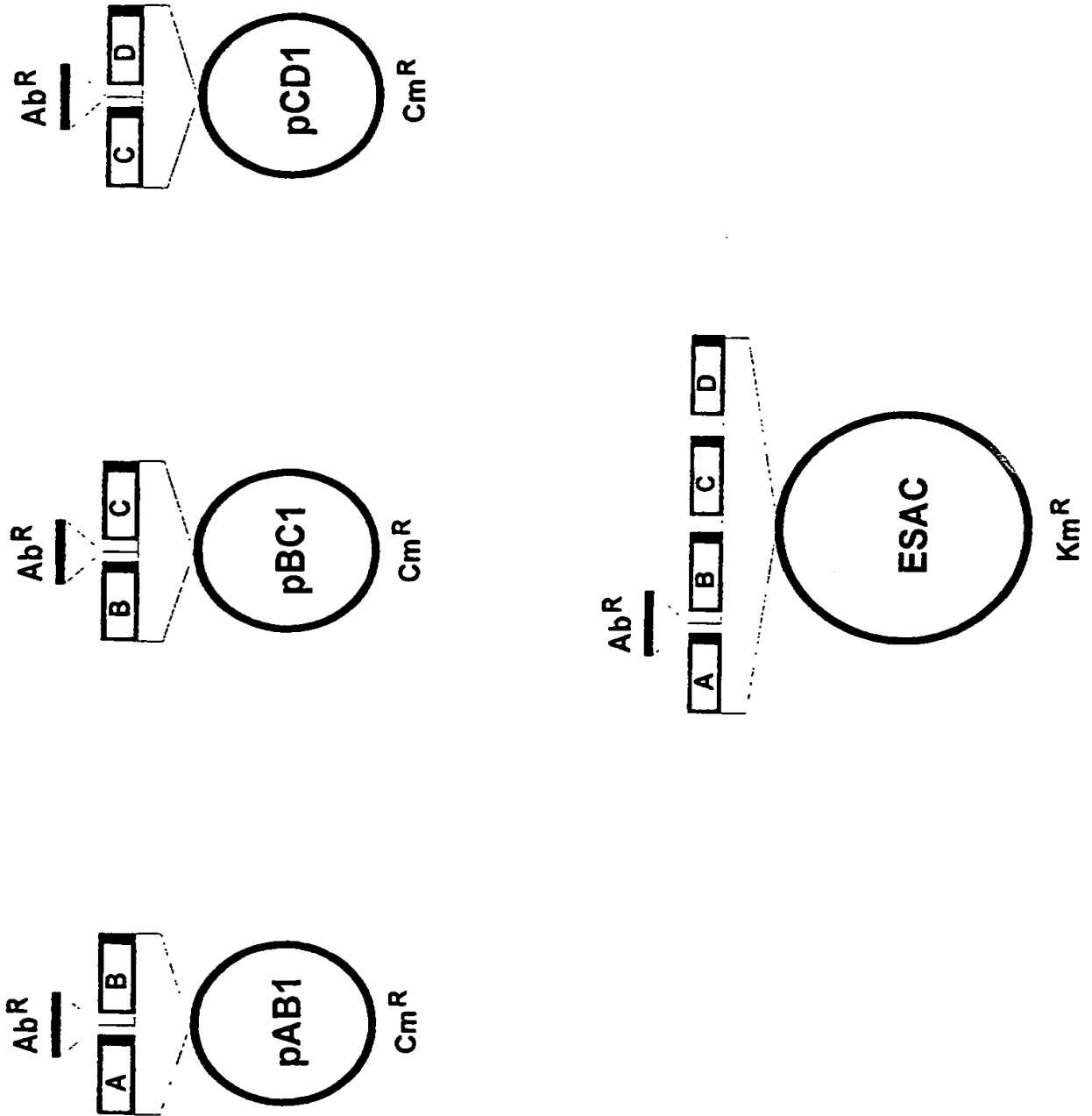


Figure 6

BSI 0003 EPN

M 23.06.98

96

7/11

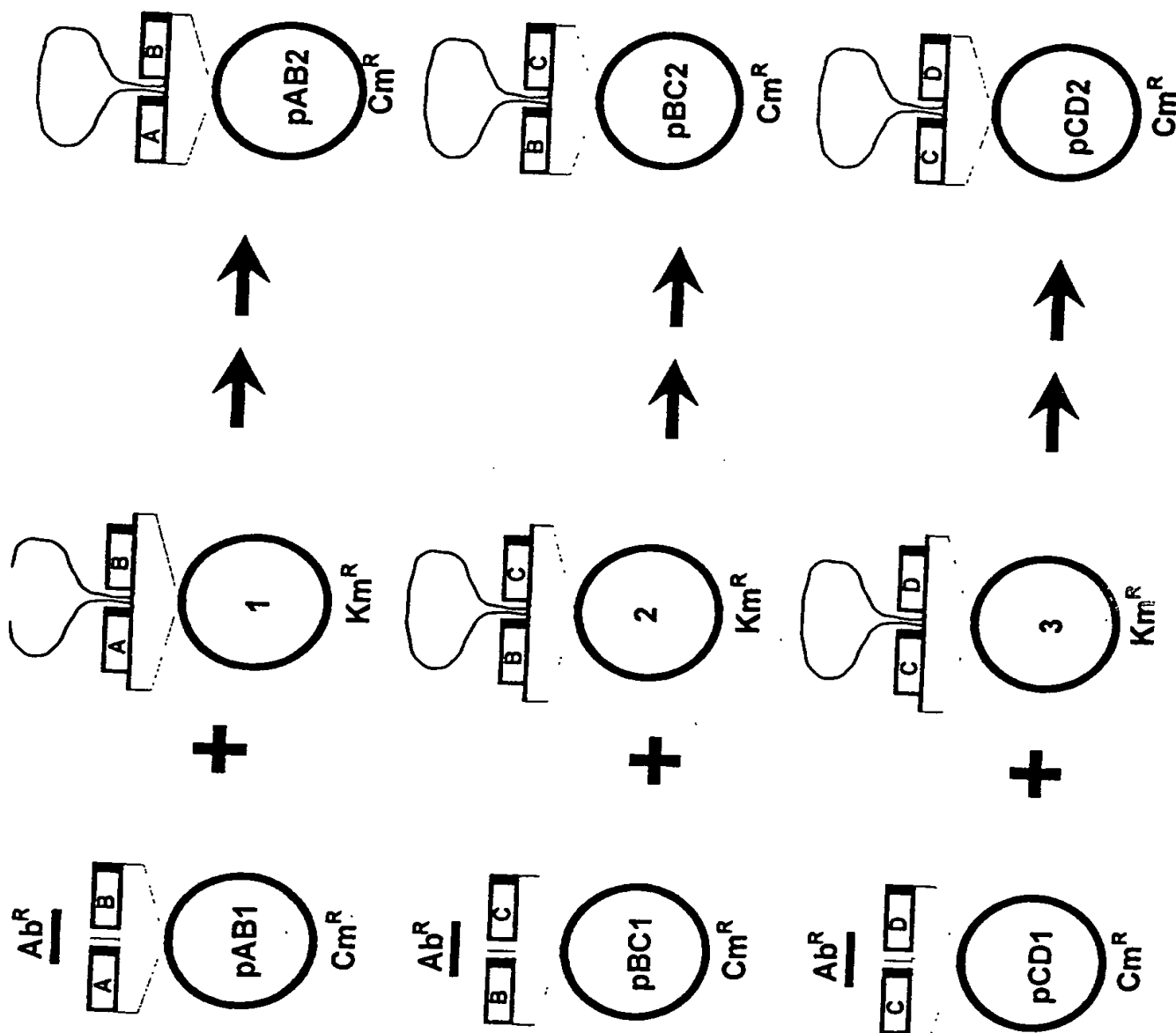


Figure 7

BSI 0003 EPN

23.08.98

97  
8/11

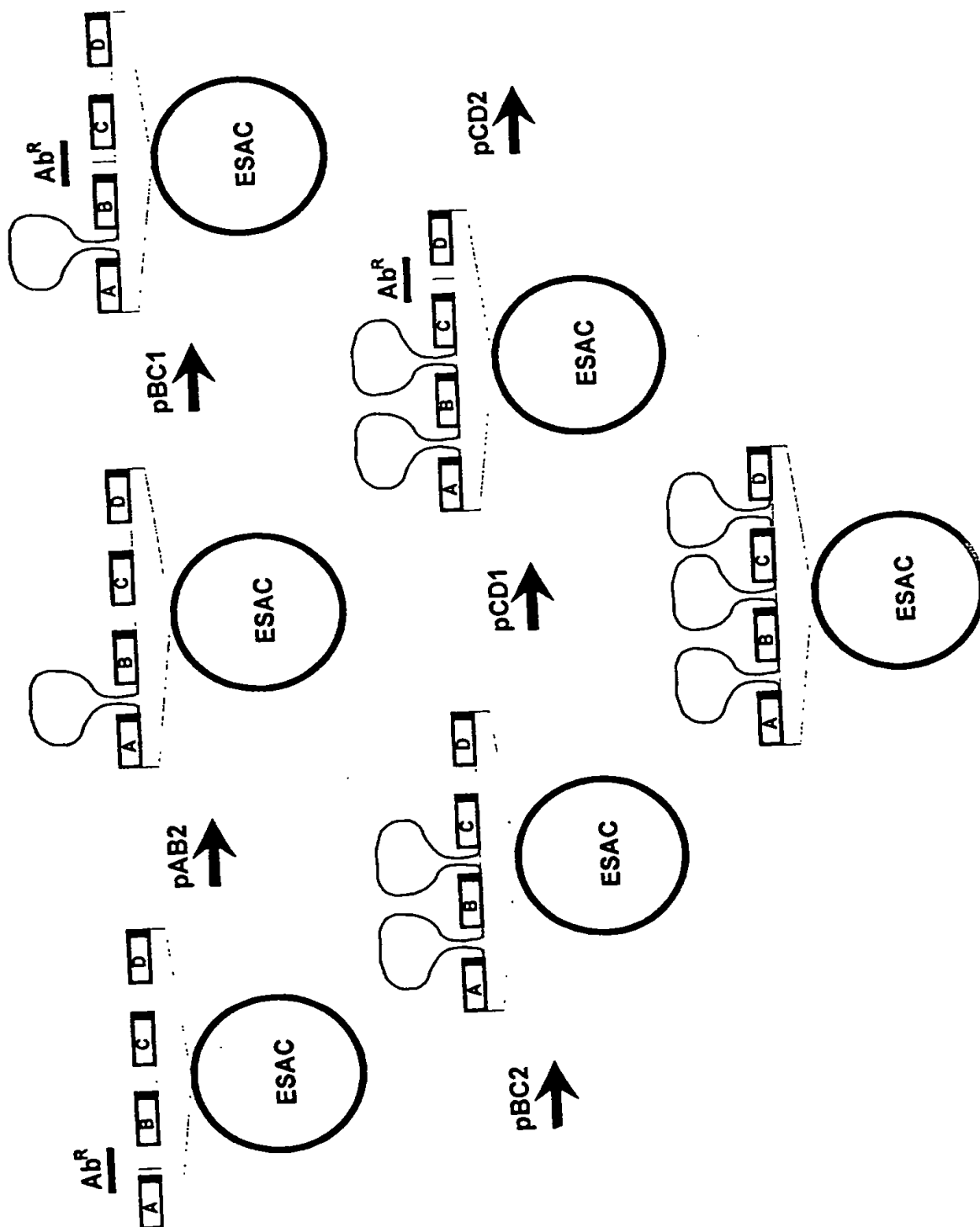


Figure 8

BSI 0003 EPN

11.23.06.98

98

9/11

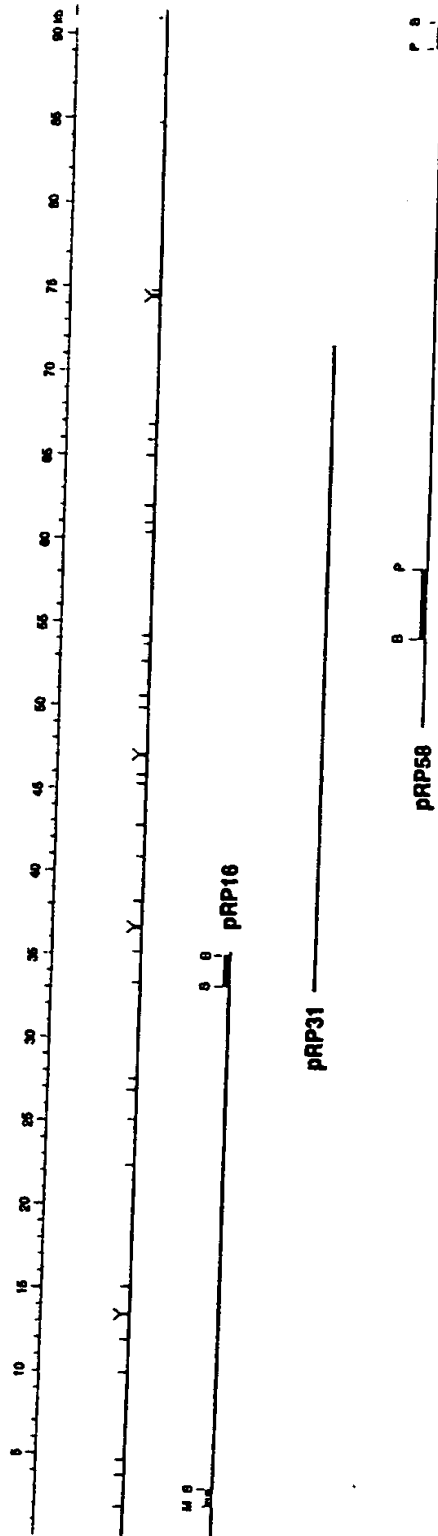


Figure 9

BSI 0003 EPN

M 23.06.98

99

10/11

**A**

1 GGATCCCGAG CACCGACCAG CCGTGGGCGG GGACGAGACA CGGGTCTCCC  
51 GGAGCCTCCC CCGACGACTC CAGCACGGCC AGGCCGCGG CCTCGACCGG  
101 GAAGCGGTAG GGCCTGTCGT CCACGGTTGA GCAGGGTGAG CAGTGCCCGG  
151 CCGGGATGGT CCGGGTCAGC CGAGGCCAGC GCGGCGGCC GGTGCTCAG

**B**

1 CCGGGAGATC CGCCGACGCC GCGGGCCGTG CACCACGGTC CTCCTGTTCC  
51 GGGCGACGGT GAACGGCAGG CAGGTCCACG GTTCCGATTT CCTGCACTTC  
101 GACGACGACG GCCTCATCGG CGAGCTCACC GTCATGGTCC GGCCGCTGTC  
151 GGGAGCGAAC GCGCTGGCCG AGGCGATGGG CGCCAGTTC GAACGGATCC

Figure 10

11 23.08.98

100  
11/11

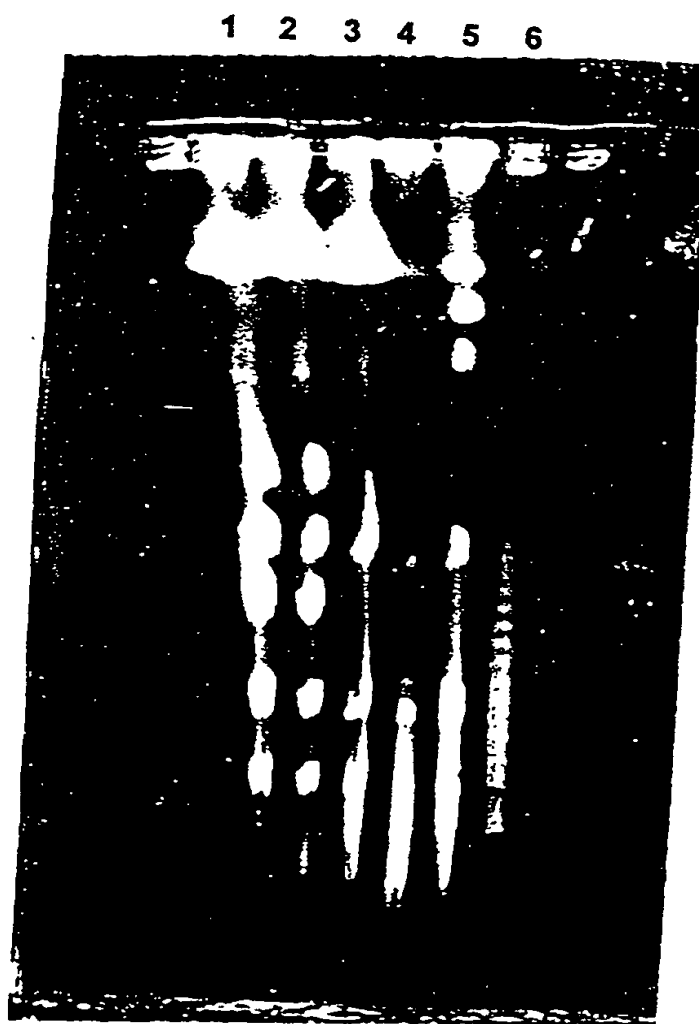


Figure 11



BSI 0003 EPN

11 23.08.98

101

## Abstract

7

The present invention provides a system for producing and modifying natural products produced by a large group of bacteria for the purpose of drug discovery, development and production. The method of the invention transfers the ability to produce a secondary metabolite from an actinomycete that is the original producer of the natural product, to a different production host that has desirable characteristics. The system involves the construction of a segment of the chromosome of the original producer in an artificial chromosome that can be stably maintained in a suitable production host. The present invention relates to recombinant DNA vectors useful for shuttling the genetic information necessary to synthesize a given natural product between a donor organism and a production host. The methods of the invention are useful in improving the yield, the purification process and for structural modification of a natural product.

